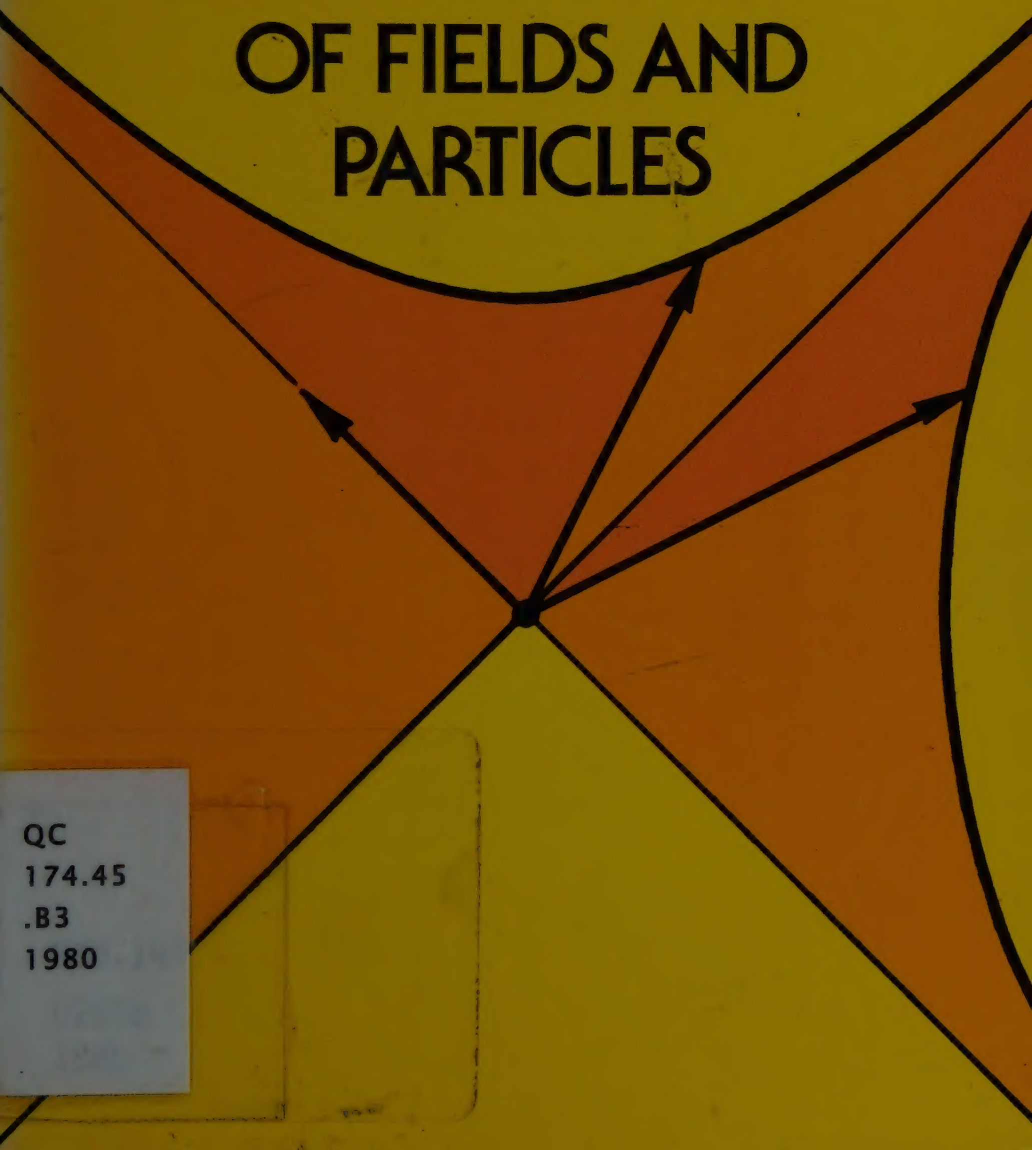
A.O. BARUT

ELECTRODYNAMICS AND CLASSICAL THEORY OF FIELDS AND PARTICLES



QC
174.45
.B3
1980

DOVER BOOKS ON PHYSICS AND CHEMISTRY

WITHDRAWN

Methods of Quantum Field Theory in Statistical Physics, A.A. Abrikosov, et al. (63228-8) [illegible]0

Electrodynamics and the Classical Theory of Fields and Particles, A. O. Barut. (64038-8) [illegible]

Electromagnetic Fields and Interactions, Richard Becker. (64290-9) $12.50

Introduction to the Theory of Relativity, Peter Gabriel Bergmann. (63282-2) $4.50

Hydrodynamic and Hydromagnetic Stability, S. Chandrasekhar. (64071-X) $11.95

Historical Studies of the Language of Chemistry, Maurice P. Crosland. (63702-6) $7.50

The Classical Electromagnetic Field, Leonard Eyges. (63947-9) $8.00

Thermodynamics, Enrico Fermi. (60361-X) $4.00

The Physical Principles of the Quantum Theory, Werner Heisenberg. (60113-7) $4.00

Atomic Spectra and Atomic Structure, Gerhard Herzberg. (60115-3) $4.00

Optics and Optical Instruments, B.K. Johnson. (60642-2) $4.50

Foundations of Potential Theory, Oliver D. Kellogg. (60144-7) $6.00

Stress Waves in Solids, H. Kolsky. (61098-5) $6.00

Hydrodynamics, Horace Lamb. (60256-7) $12.00

The Historical Background of Chemistry, Henry M. Leicester. (61053-5) $4.50

A Treatise on the Mathematical Theory of Elasticity, A.E.H. Love. (60174-9) $10.95

Fundamental Formulas of Physics, Donald H. Menzel. (60595-7, 60596-5) Two-volume set $16.00

Mathematical Physics, Donald H. Menzel. (60056-4) $7.00

Selected Papers of Physical Processes in Ionized Plasmas, Donald M. Menzel. (60060-2) $5.00

Introduction to Mathematical Fluid Dynamics, Richard E. Meyer. (61554-5) $4.50

Optical Processes in Semiconductors, Jacques I. Pankove. (60275-3) $7.50

Electromagnetic Fields and Waves, Vladimir Rojansky. (63834-0) $7.95

Elementary Mechanics of Fluids, Hunter Rouse. (63699-2) $7.00

Selected Papers on Quantum Electrodynamics, Julian Schwinger (ed.). (60444-6) $8.50

Electromagnetism, John C. Slater and Nathaniel H. Frank. (62263-0) $5.00

The Principles of Statistical Mechanics, Richard C. Tolman. (63896-0) $10.95

(*continued on back flap*)

DEO ADJUVANTE
NON
TIMENDUM

XAVIER UNIVERSITY
LIBRARY
NEW ORLEANS, LA.

CLASS...
AUTH...
ACCES...

DONATED BY

____

ELECTRODYNAMICS AND CLASSICAL THEORY OF FIELDS & PARTICLES

ELECTRODYNAMICS AND CLASSICAL THEORY OF FIELDS & PARTICLES

A.O. BARUT

Professor of Physics
University of Colorado

XAVIER UNIVERSITY LIBRARY
NEW ORLEANS, LA. 70125

DOVER PUBLICATIONS, INC.
NEW YORK

Copyright © 1964, 1980 by Asim O. Barut.
All rights reserved under Pan American and International Copyright Conventions.

Published in Canada by General Publishing Company, Ltd., 30 Lesmill Road, Don Mills, Toronto, Ontario.
Published in the United Kingdom by Constable and Company, Ltd., 10 Orange Street, London WC2H 7EG.

This Dover edition, first published in 1980, is an unabridged and corrected republication of the work originally published by The Macmillan Company in 1964.

International Standard Book Number: 0-486-64038-8
Library of Congress Catalog Card Number: 80-66841

Manufactured in the United States of America
Dover Publications, Inc.
180 Varick Street
New York, N.Y. 10014

B295e
1980

to the memory of my parents

PREFACE TO THE DOVER EDITION

This book has been out of print for many years now, but the demand for it continues. Electrodynamics and the classical theory of fields remain very much alive and continue to be the source of inspiration for much of the modern research work in new physical theories. The treatment of the foundations of electrodynamics in this book is still up-to-date. I welcome, therefore, this new Dover edition.

A number of misprints and mistakes in the original 1964 edition have been corrected. I am grateful to many colleagues and students, especially James T. Cushing, Guillermo Garcia Alcaine and W. C. Bolling, for pointing errors out to me.

A. O. B.

Boulder
September 1979

PREFACE TO THE FIRST EDITION

This book is primarily intended as an introduction on the classical level to quantized Lorentz invariant field theories. Much of the subject matter treated in this book has gained interest and been developed only after the advance of relativistic quantum theories. In spite of the shortcomings of the quantum field theory, the study of field theory and the particle-field and particle-particle action-at-a-distance interactions will be part of the education of theoretical physicists, guiding them in the foreseeable future. The need was therefore felt for a modern description of classical field theories, especially those aspects of quantum theories which one can already study on the classical level. Electrodynamics plays the central role throughout the book as the prototype of interacting fields and particles. The gravitational field is not included, but it is hoped that this book may also serve as an introduction to generally covariant field theories.

I should like to thank my colleagues W. E. Brittin, J. Currin, and D. E. Zwanziger for their suggestions after reading pertinent parts of the manuscript.

A. O. BARUT

CONTENTS

Page

Preface vii

Introduction xiii

PART I. RELATIVISTIC DESCRIPTION OF FIELDS AND PARTICLES

CHAPTER I Lorentz Transformations 3
- 1. The Physical Basis of the Lorentz Transformations 3
- 2. Mathematical Properties of the Lorentz Space 6
  - Notations 6
  - Vectors and Scalar Products 8
  - Basis 11
  - Complex Lorentz Space 11
- 3. Properties of the Lorentz Transformations 12
- 4. Special Lorentz Transformations: Applications 16
- 5. The Lorentz Group and its Representations 22
  - A. The Four Parts of the Lorentz Group 22
  - B. Correspondance with the 2 × 2 Unimodular Group 23
  - C. Spinors 25
  - D. Higher-Order Spinor and Tensor Representations 29

Page

E. Infinitesimal Generators of the Lorentz and Unimodular Groups 32
F. Complex Lorentz Group 34
6. The Principle of Relativity: Invariance and Covariance 35
7. Tensor and Spinor Fields and Momentum Space Functions 36
8. Analysis 38
9. Further Developments and Exercises 41
Bibliography for Chapter I 45

CHAPTER II Relativistic Dynamics 47
1. Proper Time Form of the Equations of Motion 48
Particles with Zero Rest Mass 52
Explicit Forms of the Minkowski Force K^{μ} 54
Angular Momentum 57
Systems of Colliding Particles 58
2. Lagrangian Form of the Equations of Motion 60
3. Canonical Form of the Equations of Motion 68
4. Electric and Magnetic Moments: Classical Spin 73
5. Further Developments and Exercises 80
Notes and Bibliography for Chapter II 83

CHAPTER III Relativistic Field Theory 85
1. Intuitive Introduction of Fields 85
2. The Electromagnetic Field 88
Basic Equations 88
Lorentz Electrodynamics 92
Covariant Form of Maxwell—Lorentz Equations 93
Gauge Transformations 94
Invariants 95
Covariant Form of Maxwell's Equations in Material Media 96
Spinor Form of the Maxwell Equations 97
3. Lagrangian Form of Field Equations 99
(A) Variations with a Fixed Boundary B
(B) Variations Involving a Change of Boundary B 103
4. Conservation Laws 105
(A) Conservation Laws in Integral Form 105
(B) Conservation Laws in Differential Form 112

Page
(C) Lagrangians Not Invariant Under Translations 115
(D) Explicit Forms of Conserved Quantities 116
5. Canonical Form of the Field Equations 119
6. Lagrangians Involving Higher-Order Derivatives 122
7. Further Developments and Exercises 127
Bibliography for Chapter III 130

PART II INTERACTIONS OF FIELDS AND PARTICLES

CHAPTER IV Equations of Motion and Their Solutions 135
1. Interactions of Fields with External "Currents" 135
2. Interactions of Fields with a Particle 138
3. Interactions Between Fields 142
4. Solutions of Field Equations: Green's Functions 148
5. Further Developments and Exercises 163
Bibliography for Chapter IV 164

CHAPTER V Radiation and Radiation Reaction 165
1. Radiation Field of a Moving Particle 165
Lienard-Wiechert Potentials 165
The Field Tensor $F^{\mu\nu}$ 168
2. Properties of the Radiation Field 170
Null Fields 170
Plane Wave Decomposition of the Radiation Field 172
Energy and Momentum of the Radiation Field 175
3. Canonical Formalism for the Radiation Field in Terms of the Transverse Vector Potential 177
4. Energy and Momentum Radiated 179
5. Radiation Reaction 184
(A) Energy Balance 184
(B) Interacting Fields and Particles 186
(C) Finite Part of Self-Force 187
(D) Mass Renormalization 190
6. Equations of Motion with Radiation Reaction 195
7. Theory of the Electromagnetic Mass 199
8. Further Developments and Exercises 203

Page

Bibliography for Chapter V 211
CHAPTER VI Action-at-a-Distance Electrodynamics 213
1. The Action Principle of Fokker-Schwarzschild-Tetrode 214
2. Action Principle with Self-Energy 217
3. Mass Renormalization 219

Related Mathematical Books 221
Author Index 225
Subject Index 231

INTRODUCTION

The two basic and most fundamental concepts of classical "atomistic" physical theories—in contrast to phenomenological theories like thermodynamics—are that of *particle* and *field*. Broadly speaking, this book deals with the interactions of systems consisting of particles and fields and, in particular as a prototype, with the interactions of charged particles with the electromagnetic field. These two concepts or entities are quite distinct in classical theories and they are described by very different mathematical apparatus: on the one hand, space-time points as the positions of the particles; on the other hand, functions of space-time as the field intensity or the potential. Only in relativistic quantum theories do we arrive at a synthesis of these two concepts, namely to attempt to describe everything in terms of fields alone, or in terms of particles alone. We cite the quantum field theory and the more recent S-matrix theory for these two directions. Even in classical theories one can find similar directions: the action-at-a-distance electrodynamics treats the interactions of particles without fields; and one can also represent in a limited sense the particles by classical fields. The process of synthesis is still unfinished and incomplete, even in the domain of quantum theories. In the main, particles and fields in this book have been treated distinctly. Many of the physical and mathematical ideas and methods of relativistic classical theories are also basic to relativistic quantum theories.

The treatment throughout the book is relativistic; we have attempted, as much as possible, a coordinate free (or covariant) form of the equations both for particles and the fields. The special theory of relativity is an outcome of Maxwell's equations and the relativistic space-time is the natural framework to use in electrodynamics and in classical field theories. For this reason we have attempted in Chapter I, Part I, to give a fairly complete discussion of space-time, Lorentz transformations, Lorentz-group and tensor and spinor fields. A fluent knowledge of the material of this chapter is essential, not only for this book but for many other branches of theoretical physics, in particular for relativistic quantum theory. However, some sections, such as the representations of the Lorentz-group, may be omitted at first reading.

The fundamental problem of dynamics is to describe the trajectories of particles for given external forces. Various relativistic forms of this problem are treated in Chapter II. In field theory the basic entities are a set of functions of space and time which describe a field. For example, the six functions $\underset{\sim}{B}(x)$, $\underset{\sim}{E}(x)$ describe the Maxwell field. These functions are not arbitrary; the class of admissible functions is defined by the solutions of the so-called field equations. The general dynamical principles to obtain the field equations and the important problem of the conservation laws are discussed in Chapter III. In electrodynamics we take mainly the microscopic point of view of Lorentz which deals with fields and charges in vacuum. The Maxwell equations in matter, i.e., in metals or dielectrics, are of phenomenological nature using average quantities, and are discussed only briefly in this chapter.

To the initial value problem of dynamics there corresponds in field theory the so-called Cauchy problem. Given the field quantities and their derivatives everywhere at a given time, the problem is to determine the field at a later time, In particular, it is of importance to find the solutions of the field equations in the presence of interactions. The interactions of fields with the particles or, more generally, with other fields, are expressed by the source terms in inhomogeneous field equations. The problem is then to find the resultant field at a space-time point, given a distribution of sources, and the "initial fields" again everywhere but at a given time. (Chapter IV.)

As the particles influence the fields by producing their own field as a result of their motion, the field itself, on the other hand, influences the motion of these sources. The first is observed as radiation and the second as radiation reaction. This feedback between the particles and the fields is discussed in Chapter V. Here also we come to questions concerning the limitations of the classical field theories and classical dynamics. For, irrespective of the problem of complete applicability of the theory to nature—quantum effects, of course, will come in before these limitations become important—the question arises whether the classical theory of interacting fields and particles is a consistent theory. The classical nonrelativistic mechanics is an example of a consistent theory; however, here the radiation and radiation reaction effects are completely neglected. With the inclusion of these effects, there are serious problems concerning the solutions of the equations of motion and the infinities occurring in the theory resulting from the mass-point character of the particles. The possible solutions of these problems are discussed.

Finally, in Chapter VI, we discuss a purely mechanical approach to the problem of interactions of charged particles, namely the so-called action-at-a-distance formulation of electrodynamics, as in the ordinary mechanics. Here the field concept is eliminated, and the equation of motion of each particle is written down taking into account "forces" resulting from all other particles.

This is then an outline and survey of the book. We have concentrated on the development of the general theory of the subject. A number of interesting applications and related subjects have been included as exercises or further reading, with brief hints and references. For the background material in electrodynamics, the recent book of J. D. Jackson, *Classical Electrodynamics* (New York: John Wiley & Sons, Inc., 1962) is recommended. For the material in this book, a more or less complete list of references is given. We remark finally that the problems of radiation and radiation reaction, and the resultant equations of motion have recently again attracted the attention of many physicists.

To keep the material continuous, alternate proofs and additional topics have been inserted in the sections "Further Developments and Exercises" at the end of each chapter. Some applications in these sections as well as some intermediary steps have been left to the reader as an exercise. The student is urged to derive these steps and do the problems, which we believe to be of great pedagogical value.

PART I

Relativistic Description of Fields and Particles

I.

LORENTZ TRANSFORMATIONS

1. *The Physical Basis of the Lorentz Transformations*[1]

The physical laws governing the behavior of fields and particles are expressed in terms of space-time coordinates $\mathbf{x}$, t and the functions of the coordinates.[2] On the other hand, the physical laws describe permanencies in nature that are independent of any coordinate frame: a physical process will take place irrespective of what coordinate frame the observer may choose from which to observe it. The arbitrariness resulting from the choice of the coordinate frames must be eliminated from the formulation of the laws. Consequently, the transformations between different possible frames of reference and the quantities which are invariant under these transformations assume a fundamental importance. The physical laws will be written in such a way that their content and form is the same for a class of observers; for an objectively identical situation only the numerical values of observed quantities may change from observer to observer.

Frames of reference in which free particles move uniformly are called *inertial frames of reference*. This concept, formally, is the same both in classical mechanics and in the special theory of relativity.[3]

---

[1] A partly deductive description of the Special Theory of Relativity is attempted here. For a historical development and references, see W. Pauli, *Theory of Relativity* (New York: Pergamon Press, 1958), Secs. 1 and 2.

[2] In contrast to this, the quantum-mechanical S-matrix theory attempts to formulate the laws of elementary particles, without space-time coordinates, in terms of directly observable quantities such as scattering amplitudes, and without the direct use of coordinates. Of course, Lorentz transformations are still used indirectly in momentum space.

[3] In the general theory of relativity even the inertial frames of reference are not singled out. These are eliminated by the use of a more general class of frames of reference, W. Pauli, *loc. cit.*

In classical mechanics the correspondence between different inertial frames is expressed by the *Galilean transformation law*: two coordinate frames moving with respect to each other with the constant relative velocity $\mathbf{w}$ are related by

$$\mathbf{x}' = O\mathbf{x} + \mathbf{w}t, \qquad t' = t, \tag{1.1}$$

where O is an arbitrary orthogonal transformation and expresses the relative position in space of the two coordinate frames at $t = 0$. The transformations, Eq. 1.1, which go hand in hand with Newton's laws, define the class of classical inertial frames. These frames are characterized therefore by the totality of the set $(O, \mathbf{w})$ with $\mathbf{w}$ any vector, O any orthogonal transformation. Newton's laws are the same for all such inertial frames although the numerical values, say, of the coordinates of the particles are different.

Other or more general physical laws (or theories) may not be valid for the class of coordinate frames defined by Eq. 1.1. A difficulty was created for physics when an attempt was made to test the validity of light phenomena for all Galilean inertial frames defined by Eq. 1.1. In the case of the propagation of the electromagnetic waves it has been established experimentally that the velocity c of the propagation is constant and the same for all inertial frames, contrary to the Galilean rule of the addition of velocities as derived from Eq. 1.1.[1] On the theoretical side, the phenomena of light propagation are correctly described by Maxwell's equations which are not valid for all Galilean inertial frames, or mathematically speaking, which are not invariant under the transformations Eq. 1.1.

This difficulty was solved by the special theory of relativity. The *Lorentz transformations* developed below will define a new class of inertial frames for electromagnetic phenomena and, because it is believed that the relations of physical phenomena to observers is the same for all physics, for all other phenomena until we are met with a new difficulty.

The study of the transformations of the type of Eq. 1.1 adds a new dimension to physical theories. On the one hand, one can study the allowed transformations for a given theory; on the other hand, one can try to construct new theories which will allow a given set of transformations. All theories are associated with a set of transformation laws that define the class of observers for which the theory is valid. Relativistic theories of mechanics, hydrodynamics, thermodynamics, statistical mechanics, and so on, are all generalizations of the corresponding nonrelativistic theories with Eq. 1.1 replaced by the Lorentz transformations.

---

[1] Recently new experiments have been performed with masers, which are about ten times more accurate than the early Michelson-Morley interferometer experiments. J. P. Cederholm, G. F. Bland, B. L. Havens, and C. H. Townes, *Phys. Rev. Letters*, *1*, (1958), 342.

After this background we go on to the development of Lorentz transformations. Because the transformation laws, Eq. 1.1, are empirically valid in the low-velocity domain, the Lorentz transformations we are seeking must reduce to the form Eq. 1.1 in the limit of small velocities. To find the proper generalization, let us consider the example of light propagation in detail.

For a light signal propagating as a spherical wave the equation of the wave front is

$$c^2t^2 - x_1^2 - x_2^2 - x_3^2 = 0.$$

The constancy of the velocity of light c implies now that in any other inertial frame to be defined, which at $t = 0$ coincided with the first, the equation of the wave front is

$$c^2t'^2 - x_1'^2 - x_2'^2 - x_3'^2 = 0,$$

i.e. again a spherical wave with the same velocity c. This is only possible if $t' \neq t$, because $x_i' \neq x_i$. Thus, the transformation law $(\mathbf{x}', t') \leftrightarrow (\mathbf{x}, t)$ must be such that

$$x_0^2 - r^2 = 0 \quad \text{implies} \quad x_0'^2 - r'^2 = 0, \quad \text{and } \textit{vice versa}, \tag{1.2}$$

where we have put $x_0 = ct$. Furthermore, the transformations must have an inverse because we can interchange the roles of the coordinates $(\mathbf{x}, t)$ and $(\mathbf{x}', t')$ and map the origin into the origin, as we have assumed the two frames to be coinciding at $t = 0$. Consider the values of x_0 and r such that $x_0^2 - r^2 = a \neq 0$ (i.e., points not on the light front). The quantity a has the dimension of (length)^2. For dimensional reasons $x_0'^2 - r'^2$ must be, therefore, a linear function of a, and by previous requirements actually equal to a.[1] The transformations which are uniquely determined by the single relation

$$x_0^2 - r^2 = x_0'^2 - r'^2 = a, \qquad a \text{ real}, \tag{1.3}$$

are then the Lorentz transformations. We shall see that they will go over to Eq. 1.1 in the limit of small velocities. Eq. 1.3 also defines the set of inertial frames for electrodynamics and all special relativistic theories.

From an experimental point of view the constancy of the velocity of light is not enough to obtain Eq. 1.3. Two other experiments must be invoked to assert this, even for linear transformations.[1] These are the Kennedy-Thorndike experiment[2] showing the independence of the transit time of light on any closed path on the velocity of the earth, and the Ives-Stilwell[3] experiment on the second-order Doppler shift, to be treated later.

---

[1] Sometimes the requirement that the transformation is linear is also introduced but this has no immediate physical meaning except that "in the small" we may consider the transformations to be (approximately) linear. The linearity will follow from Eq. 1.3.

One can also start, as did Lorentz and Poincaré, from Maxwell's electrodynamics and see under what transformations these equations are invariant. One again arrives at Eq. 1.3. In any case, whether derived inductively or deductively, we take Eq. 1.3 as the defining equation of inertial frames. Correspondingly we *define* all transformations that leave the expression

$$x^2 \equiv x_0^2 - x_1^2 - x_2^2 - x_3^2$$

invariant as *homogeneous Lorentz transformations*. More generally the *inhomogeneous Lorentz transformations* are defined to be those which leave the expression

$$(x - y)^2 \equiv (x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2$$

invariant. These differ from the homogeneous transformations by a translation of space and time coordinates. The special theory of relativity has been introduced here in coordinate space, as was the case historically. One can equally well introduce it in momentum space starting from the frequency and propagation vector of light waves (see Sec. 4.C of this chapter).

A thorough knowledge of the properties of Lorentz transformations and their representations is necessary for the following chapters. The next sections are devoted mainly to these mathematical questions.

2. *Mathematical Properties of the Lorentz Space*

Notations

We introduce a 4-dimensional vector space of *four-vectors* x with components x^0, x^1, x^2, x^3, and define a real *norm* (or *metric*),

$$x^2 \equiv x^{0^2} - x^{1^2} - x^{2^2} - x^{3^2}, \tag{1.4}$$

and a *scalar product* of two four-vectors x and y,

$$xy \equiv x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3. \tag{1.5}$$

---

[1] H. P. Robertson, *Rev. Mod. Phys.*, *21*, (1949), 378.

[2] R. J. Kennedy and E. M. Thorndike, *Phys. Rev.*, *42*, (1932), 400.

[3] H. E. Ives and C. R. Stilwell, *J. Opt. Soc. Amer.*, *28*, (1938), 215; *31*, (1941), 369.

The vectors transformed under a Lorentz transformation L will be denoted by $x' = Lx$, $y' = Ly$, $\cdots$. From the postulated invariance of the norm; i.e., $(Lx)^2 = x^2$, we can deduce the invariance of the scalar product (1.5) under Lorentz transformations. For this purpose consider the vector $x + y$. Because

$$[L(x + y)]^2 = (x + y)^2,$$

we obtain, by squaring both sides,

$$x'y' = LxLy = xy.$$

There is a one-to-one correspondence between the vectors of the four-dimensional Euclidian space R^4 and this Lorentz space L^4.[1]

Every four-vector x will have a positive definite Euclidian norm $(x, x) \geq x^{0^2} + x^{1^2} + x^{2^2} + x^{3^2}$ in R^4 and an indefinite norm, Eq. 1.4, in L^4. Eq. 1.4 can also be written within the framework of a positive definite metric as

$$x^2 = (x, Gx); \qquad xy = (x, Gy), \tag{1.4'}$$

where G is the 4×4 matrix

$$G = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} = (g_{\mu\nu}), \tag{1.6}$$

with the matrix elements $g_{00} = -g_{kk} = 1$, $g_{\mu\nu} = 0$ if $\mu \neq \nu$.[2] Thus we shall use, equivalently,

$$xy = (x, Gy) = g_{\mu\nu}x^\mu y^\nu = g^{\mu\nu}x_\mu y_\nu = x_\mu y^\mu, \tag{1.5'}$$

where the so-called *covariant components* x_μ of the coordinates are given by

$$x_\mu = g_{\mu\nu}x^\nu, \qquad x_\mu : (x^0, -x^1, -x^2, -x^3),$$

with $g^{\mu\nu}g_{\nu\sigma} = g^\mu_\sigma \equiv \delta^\mu_\sigma$. The *contravariant components* of x are x^μ.

These various notations for the norm and the scalar product are convenient and common. We must observe throughout the distinction between the covariant and the contravariant indices.

---

[1] To distinguish from the Minkowski space $(x^1, x^2, x^3, x^4 = ict)$, we denote the space with the indefinite norm (4) as the Lorentz space. To avoid the use of complex numbers, we shall use throughout the metric (1.4.)

[2] We shall use Greek indices $\mu, \nu, \cdots \sigma, \rho, \cdots$ to run from 0 to 3 and Latin indices $i, j, k, \cdots$ to run from 1 to 3, and the summation convention for both types of indices.

Vectors and Scalar Products

We divide the vectors x according to their norms into three groups:

$x^2 > 0$, time-like vectors,

$x^2 = 0$, light-like or null vectors,

$x^2 < 0$, space-like vectors.

All light-like points form the surface of a three-dimensional cone in the four-dimensional space, called the *light cone*. The light cone can be visualized as follows: a spherical light wave diminishing down to the origin (backward cone) and then expanding to infinity (forward cone), all, of course, with the velocity c. Vectors with $x^2 = a =$ constant lie on three-dimensional hyperboloids. The two-dimensional picture in Fig. 1 schematically illustrates these groups of vectors.

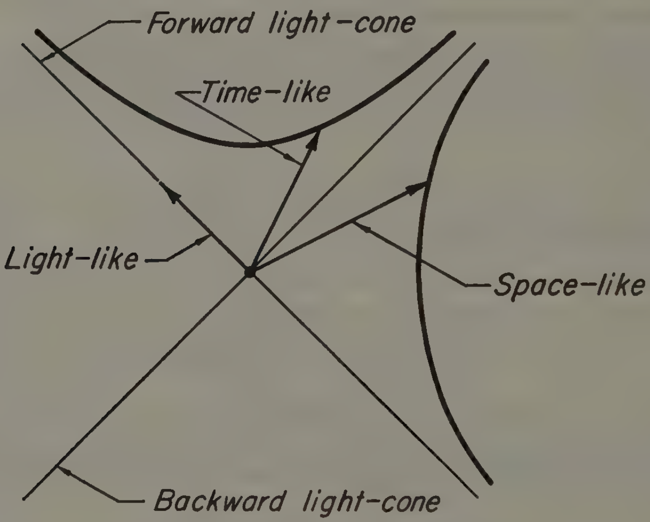


Fig. 1. Lorentz Space and three kinds of four vectors

All time-like vectors are inside the light cone and the space-like ones are outside. The slope of the line connecting any two points indicates the velocity of the signals between the points. Thus space-like points cannot be reached from each other by signals with velocities equal or less than c. This is the so-called *causality requirement* which plays an important role in relativistic theories.

The invariant scalar product $xy = (x, Gy) = x_\mu y^\mu$ may be completely defined by the following two properties:

$$xy = yx, \qquad (\alpha x + \beta y)z = \alpha xz + \beta yz\,; \qquad \alpha, \beta \text{ real numbers.} \tag{1.7}$$

We then obtain by simple algebra the following identities:

$$\begin{gathered}(x \pm y)^2 = x^2 + y^2 \pm 2xy,\\ (x + y)^2 + (x - y)^2 = 2(x^2 + y^2),\\ (x + y)(x - y) = x^2 - y^2.\end{gathered} \tag{1.8}$$

Two four-vectors x and y are called *orthogonal* if $xy = 0$. Thus every light-like vector is orthogonal to itself. It follows from the above identities that the sum and the difference of any two light-like vectors are orthogonal and have equal and opposite norms. Two orthogonal light-like vectors are at most multiples of each other, for from $xy = 0$ and $x^2 = y^2 = 0$ it follows that $\mathbf{x}\cdot\mathbf{y} = (\mathbf{x}^2)^{1/2}(\mathbf{y}^2)^{1/2}$; hence the spatial parts, and therefore the vectors, are proportional to each other. A simple way of constructing arbitrary orthogonal vectors is to take *any* two vectors of equal norm and form the sum and the difference of them.

We shall now prove some more useful relationships between the vectors.

(1) All vectors x orthogonal to a time-like vector z are space-like and form a three-dimensional space-like subspace, for from $x^0z^0 = \mathbf{x}\cdot\mathbf{z}$, we get

$$\frac{|x^0z^0|}{|z^0|} = \frac{|\mathbf{x}\cdot\mathbf{z}|}{|z^0|} < \frac{|\mathbf{x}\cdot\mathbf{z}|}{(\mathbf{z}^2)^{1/2}},$$

because z is time-like, or, squaring both sides,

$$x^{02} < \frac{|\mathbf{x}\cdot\mathbf{z}|^2}{\mathbf{z}^2} \leqq \mathbf{x}^2\,;$$

hence x is space-like. By the same argument, every vector orthogonal to a light-like vector is space-like, except the previously mentioned multiple of the light-like vector itself. There are only two orthogonal space-like vectors, both orthogonal to the light-like vector. Obviously there are three mutually orthogonal space-like vectors. This completes the orthogonality properties of the four-vectors.

(2) Analogous to the decomposition of ordinary vectors into orthogonal components, we can represent every vector x in the form

$$x = y + \tau z, \tag{1.9}$$

where z is a fixed time-like vector and y a *unique* space-like vector orthogonal to z determined by x, and τ is a real parameter. τ and y are then the components of x with respect to z. If $z:(1, \mathbf{z})$ and $y:(0, \mathbf{y})$, then τ is the time component of x. τ can be considered as an invariant time coordinate because the decomposition, Eq. 1.9, is invariant under Lorentz transformations. Similarly y will be the spatial, three-dimensional part of the four-vector z in invariant form.

$$y = x - \tau z = x - z(xz). \tag{1.9'}$$

(3) The sum of two time-like vectors z_1, z_2 is again time-like if both vectors are on the same half of the light cone (z_1^0, z_2^0 have the same sign), otherwise it is time-like or space-like, depending on whether $(z_1^0 + z_2^0)^2 > (\mathbf{z}_1 + \mathbf{z}_2)^2$, or $< (z_1^2 + z_2^2)$, or light-like.

(4) The invariant *Schwarz inequality* holds for two time-like vectors in the inverse form; i.e.,

$$(z_1 z_2)^2 \geqq z_1^2 z_2^2 ; \tag{1.10}$$

equality holds if and only if z_1 and z_2 are parallel. To show this we can assume without loss of generality z_2 to be (1000). Then ${z_1^0}^2 \geqq {z_1^0}^2 - \mathbf{z}_1^2$ which proves the statement.

For one space-like vector y and one time-like vector z we have

$$(yz)^2 > y^2 z^2. \tag{1.11}$$

Schwartz's inequality does not give anything new if one or both of the vectors are light-like. For two space-like vectors y_1, y_2 the situation is a little complicated. Taking $y_2:(0, 1, 0, 0)$, we see that for ${y_1^2}^2 + {y_1^0}^2 \leqq {y_1^0}^2 < {y_1^1}^2 + {y_1^2}^2 + {y_1^3}^2$ we have the inequality

$$(y_1 y_2)^2 \geqq y_1^2 y_2^2 ; \tag{1.12}$$

otherwise, the usual inequality[1]

$$(y_1 y_2)^2 < y_1^2 y_2^2 . \tag{1.12'}$$

(5) The form of the metric, Eq. 1.4 shows that the time and the space coordinates are on a *different* footing. The special theory of relativity says only that any time-like vector is physically as good as any other as the time coordinate, if no time reversal is involved. Therefore, we call two time-like vectors *equivalent* if $z_1 z_2 > 0$. This means that z_1 and z_2 are both in the forward or in the backward light cone. This equivalence relation is reflexive, commutative and transitive. To prove the transitivity; i.e., if $z_1 z_2 > 0$ and $z_2 z_3 > 0$, then $z_1 z_3 > 0$, we can take $z_2:(1, 0, 0, 0)$. Then $z_1^0 > 0$ and $z_3^0 > 0$ implies $z_1 z_3 > 0$. Note that the sign of the scalar product is determined by the sign of their zero components.

---

[1] W. Graeub, *Lineare Algebra* (Berlin: Springer Verlag, 1958), p. 167, always assumes Eq. 1.12′. However, one can give counterexamples that hold for (12): For example, (0,1,0,0) and $(\sqrt{3/2},1,1,0)$ satisfy Eq. 1.12, but (0,1,0,0) and (1,1,1,1) satisfy Eq. 1.12′.

Basis

In writing the norm and the scalar products in Eqs. 1.4 and 1.5, we have used the components of the vectors x, y with respect to the four basis vectors (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0) and (0, 0, 0, 1). This is not necessary. The Lorentz space can be characterized directly by the properties of the scalar product given by Eq. 1.7 and the nonpositive definiteness of the norm. Moreover, any four linearly independent four-vectors can be used as a basis. In particular, we introduce an *orthogonal basis*, a set of 4 four-vectors $e^{(\lambda)}$ ($\lambda = 0, 1, 2, 3$) such that any vector x can be written as

$$x = c^{\lambda} e^{(\lambda)}. \tag{1.13}$$

It follows from our previous considerations that one of the e^{λ} has to be time-like and the other three are then necessarily space-like. Light-like vectors cannot be used. Two or three orthogonal space-like vectors span a two- or three-dimensional space-like manifold; i.e., every vector on the manifold is space-like. Time-like and light-like manifolds are one-dimensional. Thus, in order to represent *every* vector in the form of Eq. 1.13, $e^{(\lambda)}$ must satisfy[1]

$$e_{\mu}^{(\lambda)} e^{(\sigma)\mu} = g^{\lambda\sigma}. \tag{1.14}$$

Consequently, the contravariant components of x in an arbitrary orthonormal basis $e^{(\lambda)}$ are given by

$$c^{\lambda} = x e^{(\lambda)}. \tag{1.15}$$

In the decomposition, Eq. 1.9 we can take, for example, z to be the time-like leg of the basis.

Surfaces and differential properties in Lorentz space will be discussed in Sec. 8.

*Complex Lorentz Space[2]

The complex Lorentz space has found applications in relativistic quantum theories. We add to every real four-vector x an imaginary four-vector iy and write

$$\xi = x + iy.$$

Two types of generalization of Lorentz space can now be introduced. One is the ordinary complex vector space with the *real* but nonpositive definite norm, where the scalar product is characterized by the equations

---

[1] Such a basis is denoted in the literature sometimes as "Vierbein" or "four-leg".

[2] Sections or chapters with * can be omitted at first reading without any effect in the reading of the following chapters.

$$\xi\eta = \overline{\eta\xi}\,;$$

$$\xi(\alpha\eta_1 + \beta\eta_2) = \bar{\alpha}\xi\eta_1 + \bar{\beta}\xi\eta_2\,; \quad \alpha, \beta \text{ complex numbers};$$

$$(\alpha\xi_1 + \beta\xi_2)\eta = \alpha\xi_1\eta + \beta\xi_2\eta,$$

and gives for the norm $\xi^2 = x^2 + y^2$.

In the other generalization the scalar product does not involve any complex conjugation; i.e.,

$$\xi n = \eta\xi\,; \qquad \xi(\alpha\eta) = \alpha\xi\eta\,; \qquad (\alpha\xi)\eta = \alpha\xi\eta,$$

and the norm is complex: $\xi^2 = x^2 - y^2 + 2ixy$.

Both spaces reduce in the limit $y = 0$ to Lorentz space.

3. *Properties of the Lorentz Transformations*

The homogeneous Lorentz transformations leave the scalar product xy invariant; i.e.,

$$(Lx, GLy) = (x, Gy) \equiv xy\,; \tag{1.16}$$

they are thus completely characterized by the matrix equation[1]

$$L^+GL = G. \tag{1.17}$$

For real transformations, which we are considering, $L^+ = L^T$. In the metric, Eq. 1.6 all the matrix elements of L are real.[2]

From Eq. 1.17 we find that

$$\det L = \pm 1. \tag{1.18}$$

Because the identity transformation has determinant +1, all Lorentz transformations generated continuously from the identity will have determinant +1 and are called *proper homogeneous Lorentz transformations* L_+. Transformations including reflections may have determinant −1 and will be denoted as *improper Lorentz transformations* L_-. In the notation

---

[1] A^+, A^T, A^* will denote respectively the hermitian conjugate, the transpose and the complex conjugate of the matrix A.

[2] One can also define an adjoint $\hat{L}$ to L with respect to our indefinite scalar product xy by

$$(Lx)y = x(\hat{L}y).$$

Then the Lorentz transformation are "unitary" in the sense that $\hat{L}L = I$, where $\hat{L} = G^{-1}L^+G$.

$x^{\mu\prime} = L^{\mu}_{\nu}x^{\nu}$ the matrix elements of L satisfy the following relations:

$$\begin{aligned} L^{0\,2}_{\;0} - \sum_{k=1}^{3} L^{k\,2}_{\;0} &= 1, \\ L^{0\,2}_{\;k} - \sum_{m=0}^{3} L^{m\,2}_{\;k} &= -1, \qquad k = 1, 2, 3, \\ L^{\rho}{}_{\mu} g_{\rho\sigma} L^{\sigma}{}_{\nu} &= 0, \qquad \mu \neq \nu, \end{aligned} \tag{1.19}$$

which follows from Eq. 1.17. Thus, the elements in the first column of L can be considered as components of a unit time-like vector [first line of Eq, 1.19], and those of the three other columns as components of unit space-like vectors [second line of Eq. 1.19]. These 4 four-vectors are mutually orthogonal [third line of Eq. 1.19] and form a basis, namely, that obtained from (1000), (0100), (0010), and (0001) by the application of L.

We divide the Lorentz transformations, according to the sign of $L^{0}{}_{0}$, into two groups: *orthochronous transformations* $L^{\uparrow}$ with $L^{0}{}_{0} \geqq 1$, and *antichronous transformations* $L_{\downarrow}$ with $L^{0}{}_{0} \leqq -1$. Note that $|L^{0}{}_{0}| \geqq 1$ always. Under $L^{\uparrow}$ the sign of the zero component of time-like vectors are unchanged. This is clear if the time-like vector x is (1000). In the general case we have

$$(L^{0}{}_{k}x^{k})^{2} \leqq \left(\sum_{k} L^{0\,2}_{\;k}\right)\left(\sum_{k} x^{k^{2}}\right) = (L^{0\,2}_{\;0} - 1)\sum_{k} x^{k^{2}} < L^{0\,2}_{\;0} < L^{0\,2}_{\;0}x^{0^{2}},$$

where use is made of Eq. 1.19 and the fact that x is time-like. Hence

$$|L^{0}{}_{0}x^{0}| > |L^{0}{}_{k}x^{k}|.$$

In other words, $L^{\uparrow}$ maps one time direction into an equivalent time direction; it operates within an equivalence class of time-like vectors. On the other hand, $L_{\downarrow}$ changes the sign of x^{0} and connects the two equivalent classes.

Inhomogeneous Lorentz transformations. These transformations can be written in the form

$$x' = Lx + a = \mathscr{L}x \tag{1.20}$$

where a is an arbitrary four-vector. In this case, not x^{2} but the relative norm $(x - y)^{2}$ is invariant and the origin is mapped into a. Because the inhomogeneous Lorentz transformations include the translations, they are physically more important. Any $\mathscr{L}$ is obtained as the product of a translation T and a proper Lorentz transformation L. We write symbolically

$$\mathscr{L} = (a, L). \tag{1.21}$$

The identity transformation is $(0, I)$. For the product of two inhomogeneous transformations (a_1, L_1) and (a_2, L_2); i.e.,

$$x'' = L_2 x' + a_2 = L_2(L_1 x + a_1) + a_2\,,$$

we get the rule

$$(a_2, L_2)(a_1, L_1) = (a_2 + L_2 a_1, L_2 L_1) = (a, L), \tag{1.22}$$

which can be written in the matrix form

$$\begin{pmatrix} 1 & 0 \\ a_2 & L_2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ a_1 & L_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_2 + L_2 a_1 & L_2 L_1 \end{pmatrix}. \tag{1.22$'$}$$

Thus, the inverse of (a, L) is $(-L^{-1}a, L^{-1})$.

Eigenvalues of the homogeneous Lorentz transformations. These are defined by the equation

$$Lv = \lambda v, \tag{1.23}$$

or by the roots of the equation $\det(L - \lambda I) = 0$. Let v_1 and v_2 be any two eigenvectors with eigenvalues λ_1, λ_2. Then

$$(Lv_1)(Lv_2) = v_1 v_2 = \lambda_1 \lambda_2 v_1 v_2\,; \qquad v^2 = \lambda^2 v^2.$$

We have the following possibilities: (a) If v is *not* light-like, then $\lambda^2 = 1$ or $\lambda = \pm 1$. (b) If both v_1 and v_2 are not light-like and $\lambda_1 \neq \lambda_2$, then v_1 and v_2 are orthogonal. This is also true if v_1 and/or v_2 are light-like, but with $\lambda_1 \neq \lambda_2$, and if both are different from zero. (c) If λ_1 is an eigenvalue, so is $1/\lambda_1$. The eigenvectors of complex Lorentz transformations can be discussed in a similar fashion.[1] (See Exercise 2.)

Decomposition of homogeneous Lorentz transformations into plane reflections. We define a special Lorentz transformation $\sum_n$ by

$$\sum_n x = x' = x - 2\frac{(nx)}{n^2}\,n, \qquad n^2 \neq 0, \tag{1.24}$$

for all x. n is a four-vector, and $\sum_n$ represents a reflection of the vector x through a hyperplane orthogonal to n. For example, if n:(1000), then $x^{0'} = -x^0$, $\mathbf{x}' = \mathbf{x}$; hence $\sum_n$ in this case is the time-reflection transformation. $\sum_n$ will be called a *plane reflection* or a *symmetry*. We derive the following properties of $\sum_n$ from its definition, which justify this notation and which are self-explanatory:

---

[1] E. P. Wigner, *Annals of Mathem.*, *40*, (1939), 149, p. 160.

(*a*) $\sum_n = \sum_{-n}$.

(*b*) $\sum_n n = -n$.

(*c*) $\sum_n x = x$, if $(nx) = 0$, i.e. if x is "orthogonal" to n.

(*d*) $\sum_n^2 = I$.

Furthermore, for two vectors n and $n' = Ln$, the corresponding symmetries are connected by a similarity transformation with L; i.e.,

$$L \sum_n L^{-1} = \sum_{n'}, \qquad n' = Ln. \tag{1.25}$$

To prove Eq. 1.25, we form

$$L\sum_n L^{-1}x = L\sum_n y = L\left(y - 2\frac{(ny)}{n^2}n\right) = Ly - 2\frac{(ny)}{n^2}Ln = x - 2\frac{nL^{-1}x}{n^2}Ln$$

$$= x - 2\frac{(Ln)x}{n^2}Ln = x - 2\frac{(n'x)}{n'^2}n' = \sum_{n'} x.$$

It follows from Eq. 1.25 that $\det \sum_n = \det \sum_{n'}$ and $\operatorname{tr} \sum_n = \operatorname{tr} \sum_{n'}$. By a suitable L we can always bring a time-like n into the direction n': (1000). Hence we see that

$$\det \sum_n = -1, \qquad \operatorname{tr} \sum_n = 2. \tag{1.26}$$

Therefore, if n is time-like, $\sum_n$ belongs to the $L_-^{\uparrow}$ class of Lorentz transformations; similarly, if n is space-like, $\sum_n$ belongs to $L_-^{\uparrow}$. Reflection throught a light-like vector is not defined because there is no hyperplane orthogonal to a light-like vector.

The matrix elements of $\sum_n$ are given by

$$(\sum_n)^\mu_\nu = \delta^\mu_\nu - 2\frac{n^\mu n_\nu}{n^2}. \tag{1.27}$$

In particular, we introduce four reflections $\sum_0$, $\sum_1$, $\sum_2$, $\sum_3$ of the four coordinate axes, respectively; i.e., for n: (1000), (0100), (0010), and (0001). The product $\sum_\mu \sum_\nu$ reflects two coordinates ($\mu \neq \nu$) and $\sum_\mu \sum_\nu \sum_\sigma = -\sum_\rho$, μ, ν, σ, ρ all different.

In terms of these symmetries we have the following *decomposition theorem:*[1] Any homogeneous Lorentz transformation L can be written as a product of 4 symmetries at most.

---

[1] L. Michel, *Nuovo Cimento*, *S14*, (1959), 95. This is a special case of a more general theorem: Linear transformations in an n-dimensional space leaving the nondegenerate metric $g_{ij}x^ix^j$ invariant can be written as a product of n symmetries at most.

Proof:

(a) If there exists an invariant vector u, i.e., $Lu = u$, $u^2 \neq 0$, then every vector y orthogonal to u remains orthogonal after the transformation; i.e., $yu = 0$ implies $(Ly)u = 0$. Hence L transforms the three-dimensional subspace orthogonal to u into itself and, by induction, L is the product of three plane reflections at most. One can continue this procedure down to one dimension. If at any stage there is no fixed vector x, one proceeds as follows.

(b) There exists no invariant u, but a vector y such that $Ly - y = a$ is *not* light-like. Then the product transformation $(\sum_a L)$ has a fixed vector, namely the vector y itself, for

$$\sum_a Ly = Ly - 2\frac{aLy}{a^2}a = Ly - 2\frac{(Ly-y)Ly}{(Ly-y)^2}(Ly-y)$$

$$= \frac{Ly(Ly-y)^2 - 2(Ly^2 - yLy)(Ly-y)}{(Ly-y)} = y.$$

Hence $\sum L$ is the product of three plane reflections at most, or L is, in this case, the product of four plane reflections at most.

(c) There remains only the case where no such y exists; i.e., $Ly - y$ is light-like for all y. In this case $L = I$, which is the product of two reflections.

Another decomposition of the homogeneous Lorentz transformations into rotations and special proper transformations will be discussed in the next section.

4. *Special Lorentz Transformations. Applications*

The explicit form of the following particular Lorentz transformations will be useful in actual applications.

(a) Three-dimensional orthogonal transformations[1] ($L^0_0 = 1$, $L^0_k = L^k_0 = 0$) belong to the class $L^{\uparrow}$, the rotations to $L^{\uparrow}_{+}$, and the reflections to $L^{\uparrow}_{-}$. The space-reflection transformation will be denoted by $L_P = -\sum_0$.

(b) Time reflection $L_T = \sum_0 \;\varepsilon L^{\downarrow}_{-}$.

(c) Space-time reflection $L_{PT} = L_T L_P = -I$.

---

[1] For a three-dimensional rotation around a space axis n and by an angle ϕ, we have the transformation law

$$x^{0\prime} = x^0,$$
$$\mathbf{x}' = [\cos\phi + (1 - \cos\phi)\mathbf{n}\cdot\mathbf{n} + \sin\phi\mathbf{n}\times\mathbf{x}].$$

(d) *Transformation into a moving frame of reference.* Consider a coordinate system K_1 moving with a constant velocity W along the x' direction with respect to another coordinate frame K_0. A physical point P will have the coordinates x and x' with respect to K_0 and K_1. Then $x' = L_W x$ and L_W will have the form

$$L_W = \begin{pmatrix} L^0{}_1 & L^0{}_1 & 0 & 0 \\ L^1{}_0 & L^1{}_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We have four matrix elements to determine. The defining property of the Lorentz transformations, Eq. 1.17 or 1.19, gives us three equations,

$$L^{0\,2}_{\ 0} - L^{1\,2}_{\ 0} = 1, \qquad L^{0\,2}_{\ 1} - L^{1\,2}_{\ 1} = -1, \qquad L^0{}_0 L^0{}_1 = L^1{}_0 L^1{}_1 .$$

In addition we have the kinematical condition that the origin of K_1 $(x^0, Wt, 0, 0)$ is transformed into $(x_0', 0, 0, 0)$. This means that $L^1{}_0 x^0 + L^1{}_1 (x^0/c) W = 0$ for all x^0, or

$$L^1{}_0 = -L^1{}_1 \frac{W}{c}.$$

Solving these four equations for the four unknowns, we obtain

$$L_W = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = L^{-1}_{-W}, \tag{1.28}$$

where $\gamma = 1/(1-\beta^2)^{1/2}$, and $\beta = W/c \leqq 1$. L_W belongs to the class $L^{\uparrow}_{+}$ and goes over to the Galilean transformations Eq. 1.1 for $W \ll c$, as it should. We shall discuss below some applications of this important transformation. It played an important role in the development of the theory. In particular, it occurs in the works of Lorentz and Poincaré on the invariance of Maxwell's equations and, of course, in the work of Einstein.

The columns of L_W form a basis in Lorentz space. The eigenvalues of L_W are $\lambda = [(1-\beta)/(1+\beta)]^{1/2}$ and λ^{-1}; hence the corresponding eigenvectors are light-like.

(e) *Infinitesimal inhomogeneous Lorentz transformations.* These are proper Lorentz transformations differing infinitesimally from the identity,

$$x' = x + \omega x + \varepsilon, \tag{1.29}$$

where ω is an infinitesimal matrix and ε an infinitesimal four-vector. The condition Eq. 1.17 implies for the homogeneous part

$$\omega^+ G + G\omega = 0, \tag{1.30}$$

or

$$\omega_{\mu\nu} = -\omega_{\nu\mu}, \qquad \text{where } \omega_{\mu\nu} = g_{\mu\sigma}\omega^{\sigma}{}_{\nu}\,. \tag{1.30$'$}$$

Thus the infinitesimal homogeneous transformations are determined by six parameters,[1] the inhomogeneous ones by ten parameters.

Next we shall discuss some *physical applications* of the transformation L_W.

(*A*) *Relativistic Addition of Velocities.* Consider a coordinate frame K_1 moving with respect to K_0 with a velocity W_1 and a frame K_2 moving with respect to K_1 with a velocity W_2. What is the relative velocity W of K_2 with respect to K_0? Let the corresponding Lorentz transformations be denoted by L_1, L_2 and L, respectively. Then we have $Lx = L_2(L_1 x)$ for all x, or

$$L = L_2 L_1\,,$$

where each of the L, L_1, L_2 is given by Eq. 1.28 with the parameters γ, γ_1, and γ_2 respectively. A matrix multiplication and comparison of the matrix elements of both sides gives, after some calculation,

$$\beta = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}\,. \tag{1.31}$$

(*B*) *Decomposition of proper homogeneous Lorentz transformations.* We want to decompose any homogeneous transformation L into a product of rotations and a L_W. Because L connects frames that are, at most, moving uniformly with respect to each other, we can first bring the direction of the relative motion into, say, the x^1 axis by a rotation, and then apply L_W. The two resulting frames can then differ only by another rotation. The effect of these three operations is the same as that of L. Hence every L can be written in the form

$$L = R_2 L_W R_1\,, \tag{1.32}$$

where R_2 and R_1 are three-dimensional rotations. In this decomposition

---

[1] In general the number of parameters for n-dimensional orthogonal or pseudo-orthogonal transformations is $n(n - 1)/2$.

L_W is *unique* and is given by Eq. 1.28, but R_2 and R_1 are not unique. Eq. 1.32 can also be written as[1]

$$L = L'_W R'_1 . \tag{1.32'}$$

As an example, we reconsider the problem of addition of velocities for the case where K_2 is moving in a direction that makes an angle α with the x^1 axis. Let R be the rotation by an angle α, so that the two x^1 axes coincide after this rotation. Then $LRx = L_2 R L_1$, and, by comparing the $^0_0$ components of both sides, we find

$$\beta^2 = 1 - \frac{(1-\beta_1^2)(1-\beta_2^2)}{(1+\beta_1\beta_2\cos\alpha)^2}, \tag{1.33}$$

which reduces to Eq. 1.31 for $\alpha = 0$. Note that we cannot compare other matrix elements, other than $^0_0$, because we have introduced only one rotation rather than two; the x^2 and x^3 coordinates may still be rotated with respect to each other. As another application of Eq. 1.32, we derive the form of $L_{\mathbf{w}}$ connecting two frames moving with respect to each other in an arbitrary direction with velocity $\mathbf{w}$ and $\boldsymbol{\beta} = \mathbf{w}/c$. We have

$$L_{\mathbf{w}} = R_1 L_W R_2 ,$$

where

$$R_1 = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R} \end{pmatrix}, \qquad R_2 = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{S} \end{pmatrix}.$$

R and **S** are three-dimensional rotations, **R** carrying the x^1 axis into the direction of $\mathbf{w}$; **S** is a rotation around $\mathbf{w}$ to make the x^2 and x^3 axes parallel. We get

$$L_{\mathbf{w}} = \begin{pmatrix} \gamma & -\beta_1\gamma & -\beta_2\gamma & -\beta_3\gamma \\ -\beta_1\gamma & 1 + A\beta_1\beta & A\beta_1\beta_2 & A\beta_1\beta_3 \\ -\beta_2\gamma & A\beta_2\beta_1 & 1 + A\beta_2\beta_2 & A\beta_2\beta_3 \\ -\beta_3\gamma & A\beta_3\beta_1 & A\beta_3\beta_2 & 1 + A\beta_3\beta_3 \end{pmatrix}, \tag{1.34}$$

where $A = (\gamma - 1)/\beta^2$. Note that $L_{\mathbf{w}} = -L^{-1}_{-\mathbf{w}}$, as it should. It is sometimes convenient to put in Eq. 1.34

$$\gamma = \cosh X, \qquad -\boldsymbol{\beta}\gamma = \mathbf{n} \sinh X,$$

then

$$x^{0\prime} = x^0 \cosh X + \mathbf{x}\cdot\mathbf{n} \sinh X,$$

$$x' = x^0 \mathbf{n} \sinh X + [\mathbf{x} - \mathbf{n}(\mathbf{x}\cdot\mathbf{n})] + \mathbf{n}\mathbf{x}\cdot\mathbf{n} \cosh X.$$

---

[1] This follows, for example, from the polar decomposition of 2×2 unimodular matrices A which are homomorphic to restricted Lorentz transformations. See Section 5B and Exercise 18.

(*C*) *Doppler effect and abberation.* So far we have considered the transformations of the coordinates x only. Physical equations involve other quantities beside the positions of the mass points in space-time. These quantities must also transform between inertial frames to guarantee the invariance of the physical laws. Consider an electromagnetic plane wave of the form

$$\phi = Ae^{i(\omega t - \mathbf{k}\cdot\mathbf{r})},$$

where ϕ is an electric or magnetic field, or the electromagnetic potential, ω is the angular frequency, and $\mathbf{k}$ is the propagation vector. All phenomena, such as the vanishing of a light signal, or light reaching a maximum, etc., must be the same for all inertial frames. These processes are determined by the phase $\alpha = (\omega/c)ct - \mathbf{k}\cdot\mathbf{r}$ of the wave. Thus, α must be an invariant quantity under the Lorentz transformations. This suggests the introduction of a four-vector $k : (\omega/c, \mathbf{k})$ so that α is given by the invariant scalar product

$$\alpha = kx = (k, Gx). \tag{1.35}$$

All that we have said about the Lorentz space of x vectors will apply also to the space of the four-vectors $\mathbf{k}$. In particular, from its construction, we see that k is a light-like vector: $k^2 = (\omega^2/c^2) - \mathbf{k}^2 = 0$. As a matter of fact, the name "light-like" originates here; k is the propagation vector of the light.

Let us look now at our plane wave from a different moving frame of reference. The propagation vector in this frame will be

$$k' = L_{\mathbf{w}}k. \tag{1.36}$$

If the coordinate frame is moving, say, in the x^1 direction, we can use Eq. 1.28 to write Eq. 1.36 in component form. If α is the angle between $\mathbf{k}$ and the x^1 direction and α^1 the angle between $\mathbf{k}'$ and the x^1 direction, we obtain

$$\begin{aligned} \omega' &= \omega \frac{1 - \beta\cos\alpha}{(1-\beta^2)^{1/2}}, \\ \cos\alpha &= \frac{\beta + \cos\alpha'}{1 + \beta\cos\alpha'}. \end{aligned} \tag{1.36'}$$

The first equation gives the change in frequency, the second gives the change in direction of the wave as observed from a moving coordinate frame.

Introduction of the special theory of relativity in momentum space. Directly observable quantities of an electromagnetic plane wave are ω and k, defined above, satisfying

$$\omega/c = (k_1^2 + k_2^2 + k_3^2)^{1/2}.$$

Starting from these quantities and noting that a plane wave is a plane wave in every frame of reference, we arrive at the analogues of Eq. 1.2; viz.,

$$(\omega/c)^2 - \mathbf{k}^2 = 0 \text{ implies } (\omega'/c)^2 - \mathbf{k}'^2 = 0, \text{ and } \textit{vice versa}.$$

Again, for dimensional reasons, the equation $(\omega/c)^2 - \mathbf{k}^2 = a^2$ must imply $(\omega'/c)^2 - \mathbf{k}'^2 = a^2$. That this generalization not only determines uniquely the Lorentz transformations, but also has physical significance, is seen from the existence of plane waves (de Broglie waves) corresponding to particles with nonzero rest mass. Compare the discussion in Chapter II, Sec. 1.

(*D*) *Transformation of finite lengths and time intervals.* We discuss now the appearance of lengths and time intervals as seen from a moving frame. (Fitzgerald contraction and time dilatation). Consider two points P_1, P_2 in space-time[1] with coordinates x_1, x_2 in one frame, and x_1', x_2' in another frame moving relative to the first. Then

$$x_2' - x_1' = L_{\mathbf{w}}(x_2 - x_1),$$

or, for a motion in the x^1 direction, using Eq. 1.28,

$$c(t_2' - t_1') = \gamma c(t_2 - t_1) - \beta\gamma(x_2 - x_1),$$
$$x^{2'} - x^{1'} = -\beta\gamma c(t_2 - t_1) + \gamma(x^2 - x^1).$$

A length measurement in any frame is defined as the simultaneous identification of the endpoints of the interval. Thus, setting $t_2' - t_1' = 0$ and $x^2 - x^1 = l$, we get

$$l' = \frac{1}{\gamma} l = (1 - \beta^2)^{1/2} l, \qquad \textbf{(1.37)}$$

i.e., in units of the first system, the length in the second system is shortened by a factor $(1 - \beta^2)^{1/2}$. Similarly, for the measurement of a time interval at a given point, $x^{2'} = x^{1'}$, and we get

$$\Delta\tau' = (1 - \beta^2)^{1/2} \Delta\tau, \qquad \textbf{(1.38)}$$

---

[1] Points in four-dimensional space will also be called *events*.

i.e., time lags behind in the moving frame relative to the rest frame, or a process of a certain duration will take a much longer time in a moving frame as compared to a rest frame. For example, the decay time of an unstable particle when it is moving is longer than when it decays at rest.[1] The observation of Eq. 1.37 may be more difficult for the following reason: the eye records photons arriving at the retina simultaneously from all parts of the object. However, these photons have not been emitted simultaneously by all points because of the motion of the object. This effect exactly cancels the Fitzgerald contraction for objects that subtend a small solid angle, and moving circular objects will still have an approximately circular appearance.[2] For this and the much-discussed twin or clock paradox, see Exercises 1.4 and 1.5.

5. *The Lorentz Group and Its Representations*[3]

(A) The Four Parts of the Lorentz Group

The set of all Lorentz transformations form a group. This means that, given two transformations, their inverses and their product are again Lorentz transformations. The identity transformation is included in this set. The *inhomogeneous Lorentz group* $\mathscr{L}$ is characterized by the set of translations and homogeneous Lorentz transformations[4] $[(a, L)]$, and the multiplication law Eq. 1.22 or Eq. 1.22′ shows this group property. $\mathscr{L}$ has the subgroup L, the *full homogeneous Lorentz group*. The elements of L are explicitly given by the 4×4 matrices L satisfying Eq. 1.17; i.e., $L^+GL = G$. Also, on the basis of this equation, one may easily verify that L^{-1} exists and also satisfies Eq. 1.17, and, together with L_1 and L_2, the product L_1L_2 is also a Lorentz transformation. In addition, L^+ and $L^{+^{-1}}$ also satisfy Eq. 1.17.

By Eq. 1.18 the homogeneous transformations are divided into two parts: proper transformations with det $L = +1$ and the improper transformations with det $L = -1$. The proper Lorentz transformations L_+ form a subgroup, the improper ones do not. This is a consequence of the following propositions: det $I = +1$, det $L^{-1} =$ det L, and det $L_1 = +1$, det $L_2 = +1$ imply $\det(L_1L_2) = +1$.

---

[1] This has actually been observed: Compare the lifetime of, say, π mesons at rest (2.56×10^{-8} sec) and π mesons moving at a velocity $v \approx 0.75c$ [3.8×10^{-8} sec, R. P. Durbin, H. H. Loar, and W. W. Havens, Jr., *Phys. Rev. 88*, (1952), 179].

[2] R. Penrose, *Proc. Camb. Phil. Soc.*, *55*, (1959), 137; J. Terrell, *Phys. Rev.*, *16*, (1959), 1041; V. Weisskopf, *Phys. Today*, *13*, (1960), 24; *Lectures in Theoretical Physics* (New York: Interscience Publishers, 1960), Vol. III.

[3] No detailed knowledge of group theory is required for this section. Although some of the details of this section may be passed over on first reading, the essential ideas will be important in the discussions of tensors, spinors, etc.

[4] We shall denote groups and transformations by the same letter: the group $\mathscr{L}$ has elements $\mathscr{L}$, etc.

Furthermore, the orthochronous Lorentz transformations $L^\uparrow$ with $L^0_1 \geqq 1$ form a subgroup of L, but the antichronous transformations $L_\downarrow$ do not. Again we have to verify that the inverse of an element of $L^\uparrow$ belongs to $L^\uparrow$ and the product of two orthochronous Lorentz transformations is again orthochronous. Let us prove the second statement first. Let L and M be transformations belonging to $L^\uparrow$; then

$$(LM)^0{}_0 = L^0{}_\mu M^\mu{}_0 = L^0{}_0 M^0{}_0 + L^0{}_k M^k{}_0,$$

and by the Schwartz inequality for the spatial parts $L^0{}_k$, $M^k{}_0$,

$$(LM)^0{}_0 \geqq L^0{}_0 M^0{}_0 - \Big[\sum_k (L^0{}_k)^2\Big]^{1/2}\Big[\sum_k (M^k{}_0)^2\Big]^{1/2} \geqq 1;$$

the last inequality follows from Eq. 1.19. Moreover, we see that if $L^0{}_0 \geqq 1$ and $(LM)^0{}_0 \geqq 1$, then $M^0{}_0 \geqq 1$; thus the inverse L^{-1} belongs also to $L^\uparrow$ together with L. This completes the proof of the group property *of the orthochronous Lorentz group* $L^\uparrow$.

The intersection of $L^\uparrow$ and L_+ is the *restricted Lorentz group* $L^\uparrow_+$. Therefore we obtain the following four parts of the homogeneous Lorentz group: $L^\uparrow_+$, $L^\downarrow_+$, $L^\uparrow_-$, $L^\downarrow_-$. These four parts are disconnected; that is to say, two group elements in different parts cannot be connected *continuously* with each other by varying the parameters of the elements. The inhomogeneous Lorentz group $\mathscr{L}$ is a ten-parameter group; i.e., to specify a general transformation, we need to specify ten real parameters. The homogeneous group L has six parameters.[1] We see these most easily from the infinitesimal transformations, Eqs. 1.29, 1.30.

(B) Correspondence with the 2 × 2 Unimodular Group

We discuss now a very close correspondence (a homomorphism) between the restricted Lorentz group $L^\uparrow_+$ and the unimodular group of 2 × 2 complex matrices.[2] This will allow us to introduce the basic entities called spinors. Both groups are six-parameter groups, and the correspondence we are going to discuss is a generalization of the well-known homomorphism between the three-dimensional rotation group and the two-dimensional unimodular-unitary group.[1]

---

[1] The groups of transformations depending on a finite number of parameters are called *Lie groups*. Other Lie groups besides the Lorentz group or the group of rotations have found important applications in theoretical physics. In contrast to this, groups like the gauge groups of electrodynamics (see later) depend on arbitrary functions and may be called *infinite groups*.

[2] This group is called in literature $SL(2,C)$.

More precisely, we want to show that if A is a 2×2 unimodular matrix, then to $\pm A$ there corresponds a Lorentz transformation $L(A)$,

$$\pm A \to L(A) \,\varepsilon\, L_+^\uparrow, \tag{1.39}$$

and this correspondence preserves the multiplication law (i.e., it is a homomorphism in mathematical language):

$$\pm(AB) \to L(A)L(B) = L(AB). \tag{1.39$'$}$$

The homomorphism is two-to-one; both $+A$ and $-A$ correspond to the same Lorentz transformation. We shall show this important correspondence explicitly in the following form:

Let X be *any* 2×2 Hermitian matrix. Any 2×2 matrix can be written as a linear combination of identity and three basic 2×2 matrices.[2] Let

$$X = x^0 I + x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3 = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix},$$

where the coefficients x^0, x^1, x^2, x^3 are real numbers and the basic matrices σ have been chosen to be the Hermitian Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.40}$$

Writing $\sigma_0 = I$, we shall use the short notation

$$X = \sigma_\mu x^\mu.$$

Now let A be an arbitrary 2×2 complex unimodular matrix; i.e.,

$$\det A = 1.$$

Then $X' = AXA^+$ is again a hermitian 2×2 matrix and must be of the form

$$X' = x'^\mu \sigma_\mu,$$

Now $\det X' = \det X$, or explicitly

$$x^{0^2} - x^{1^2} - x^{2^2} - x^{3^2} = x^{0'^2} - x^{1'^2} - x^{2'^2} - x^{3'^2}.$$

Therefore, the matrix transformation $X' = AXA^+$ induces a transformation on the coefficient x^μ, which is a Lorentz transformation $L(A)x' = L(A)x$. To find it explicitly, we note that

$$L^\mu{}_\nu x^\nu \sigma_\mu = Ax^\mu\sigma_\mu A^+ = x^\mu A\sigma_\mu A^+,$$

---

[1] See, for example, H. Goldstein, *Classical mechanics* (New York: Addison-Wesley, 1959), p. 110.

[2] See, for example, H. Goldstein, *Ibid.*

or, comparing the coefficients of x^μ,

$$L^\mu{}_\nu \sigma_\mu = A\sigma_\nu A^+.$$

To obtain $L^\mu{}_\nu$, we multiply both sides with σ_ρ, take the trace,[1] and we get finally

$$[L(A)]^\mu{}_\nu = \tfrac{1}{2}\mathrm{tr}(\sigma_\mu A\sigma_\nu A^+) = \tfrac{1}{2}\mathrm{tr}(\tilde{\sigma}^\mu A\sigma_\nu A^+) \tag{1.41}$$

This is the desired correspondence. Clearly the matrices A and $-A$ generate the same $L(A)^\mu{}_\nu$. We have only to show that $L(A)$ defined by Eq. 1.41 belongs to the restricted Lorentz group $L_+^\uparrow$. We see this as follows: Take first the time-like vector (1000), then $X = I$ and $X' = AA^+$ contain positive definite diagonal elements, which means that $x^{0'} \geqq x^{3'} \geqq 0$; i.e., L is orthochronous. Both $A = \pm 1$ correspond to $L = I$, and because infinitesimal A's correspond to infinitesimal L's, det L must be equal to one. Moreover, there is no A corresponding to space-reflections. For example, the vector X: (0001) and a space-reflection lead to the equation $A\sigma_3 A^+ = -\sigma_3$, which is a contradiction. The fact that det $L = +1$ can also be proved by formal manipulation starting from Eq. 1.41. (See Exercise 6.)

(C) Spinors

We introduce now the concept of *group representations*. Let us consider a group G with elements $(a, b, c, \cdots)$. If to every group element a there corresponds a transformation (or matrix) D(a) in some linear vector space,

---

[1] In this calculation and others we make use of the following properties of the matrices Eq. 1.40:

$$\mathrm{tr}\,\sigma_0 = 2, \quad \mathrm{tr}\,\sigma_k = 0, \quad k = 1, 2, 3,$$

$$\sigma_j\sigma_k = i\sigma_l; \quad j, k, l \text{ cyclic}, \quad \text{or} \quad \sigma_j\sigma_k = \delta_{jk} + i\varepsilon_{jkl}\sigma_l,$$

$$\sigma_\mu{}^+ = \sigma_\mu = \sigma_\mu{}^{-1},$$

$$\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i\boldsymbol{\sigma},$$

$$\sigma_j\sigma_k\sigma_l = \sigma_j\delta_{kl} - \sigma_k\delta_{jl} + \sigma_l\delta_{jk} + i\varepsilon_{jkl},$$

or

$$\sigma_\mu\sigma_\nu - \frac{i}{2}\varepsilon_{\mu\nu\beta\alpha}\sigma^\beta\sigma^\alpha = \delta_{\mu\nu}; \quad \sigma^\mu\colon (I, -\boldsymbol{\sigma}); \; \sigma_\mu\colon (I, \boldsymbol{\sigma}).$$

The quantities ε_{jkl}, $\varepsilon_{\mu\nu\alpha\beta}$ are totally antisymmetric indices in three and four dimensions.

$$\mathrm{tr}(\sigma_j\sigma_k) = 2\delta_{jk},$$

$$\mathrm{tr}(\sigma_j\sigma_k\sigma_l) = 2i\varepsilon_{jkl},$$

$$\mathrm{tr}(\sigma_j\sigma_k\sigma_l\sigma_m) = 2(\delta_{jk}\delta_{lm} - \delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}).$$

It is also convenient, for sake of covariance, to introduce the matrices

$$\tilde{\sigma}_\mu = \sigma^\mu \quad \text{or} \quad \tilde{\sigma}^\mu = \sigma_\mu$$

such that

$$\mathrm{tr}(\tilde{\sigma}^\mu\sigma_\nu) = 2g^\mu{}_\nu.$$

called the *representation space*, in such a way that to the product (ab) of the group elements there corresponds the product of the transformations, $ab \leftrightarrow D(a)D(b) = D(ab)$, then the set of transformation $\{D(a)\}$ is called a *representation* of the group G. The representation is *faithful* if to distinct elements a and b there corresponds distinct matrices $D(a)$ and $D(b)$. If, in addition, we require $D(1) = I$, then $D(a^{-1}) = D^{-1}(a)$, and we have a representation in which all matrices are nonsingular.[1]

With this concept we see that the 2×2 unimodular matrices A discussed in the previous section form a *two-dimensional representation* of the restricted Lorentz group $L_+^\uparrow$. We already know that $A(1) = \pm I$. The equation $A(L_1L_2) = A(L_1)A(L_2)$ is an immediate consequence of the two equations $X' = A(L_1)XA^+(L_1)$ and $X'' = A(L_2)X'A^+(L_2)$. This two-dimensional representation is faithful, but *two-valued.* That is, to every group element there correspond two matrices $\pm A$.

Let us now consider the vectors in the two-dimensional representation space and denote them by

$$\xi = \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}, \tag{1.42}$$

where ξ^1, ξ^2 are the complex components of the two-vector ξ. Under the transformation $A(L)$, ξ transforms as

$$\xi' = A(L)\xi, \qquad \det A(L) = +1. \tag{1.43}$$

Two-component quantities that, under Lorentz transformation, transform according to Eq. 1.43 are called *spinors*. This is a generalization of the well-known spinors corresponding to the three-dimensional rotations R, where the matrices $A(R)$ are *unitary* in addition to being unimodular.

Up to this section we have considered vectors a as the geometrical objects that transform under Lorentz transformations according to the law $a' = La$, such as the coordinates x or momenta k. Spinors are new geometrical objects with a quite different, but definite, transformation property. The concept of group representations allows us to give a general characterization of all such objects. If $D(L)$ is any finite- or infinite-dimensional representation of the Lorentz group, then we will have geometrical objects ζ (or *tensors* in the terminology of E. Cartan) transforming under Lorentz transformations according to

$$\zeta' = D(L)\zeta. \tag{1.44}$$

ζ will have as many components or indices as the dimensionality or indices of D. In the particular case of vectors, L itself is, of course, a 4×4 representation of the Lorentz group. Other realizations of Eq. 1.44 beside Eq. 1.43 will be discussed in the following sections.

---

[1] In quantum theories it is necessary to consider more generally *representations up to a factor*, $ab \leftrightarrow \omega(a, b)D(a)D(b)$, where $\omega(a, b)$ is a phase factor.

In the case of vectors, we know that the scalar product (x, Gy) is invariant under Lorentz transformations. We look now for an invariant of the form

$$(\xi, C\eta) = \xi^{\alpha} C_{\alpha\beta} \eta^{\beta} \tag{1.45}$$

in two-dimensional *spinor space* (the indices α, β, $\cdots$ are summed over 1, 2). Because ξ and η transform according to Eq. 1.43, the invariance of Eq. 1.45 implies that C must satisfy the equation[1]

$$A^T C A = C, \qquad \det A = +1. \tag{1.46}$$

Any antisymmetric matrix C is a solution of Eq. 1.46. For if C is antisymmetric, then $A^T C A$ is:

$$A_{\delta\alpha} C^{\delta\gamma} A_{\gamma\beta} = -A_{\delta\alpha} C^{\gamma\delta} A_{\gamma\beta} = -A_{\delta\beta} C^{\delta\gamma} A_{\gamma\alpha}\,;$$

in the last step we have interchanged the dummy indices δ and γ. We take a standard form of C, namely

$$C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma_2 = -C^{-1} = C^* = -C^T, \tag{1.47}$$

and, in analogy to G in Eq. 1.6, we call it the *metric spinor*.[2]

We defined covariant components of vectors as $x_\mu = g_{\mu\nu} x^\nu$; similarly we introduce covariant components of spinors:

$$\xi_\alpha = C_{\alpha\beta} \xi^\beta, \tag{1.48}$$

so that the invariant bilinear form is

$$(\xi, C\eta) = \xi^\alpha \eta_\alpha = -\,\xi_\alpha \eta^\alpha. \tag{1.45$'$}$$

If $\xi'^\alpha = A(L)^\alpha{}_\beta \xi^\beta$, then the covariant spinors transform according to.[3]

$$\xi'_\alpha = [A^{T^{-1}}(L)]_\alpha{}^\beta \xi_\beta \equiv A_\alpha{}^\beta \xi_\beta\,. \tag{1.49}$$

---

[1] Note that although ξ and η are complex vectors, the bilinear form Eq. 1.45 is formed like that of a real vector space. The form Eq. 1.45 itself is a complex quantity. There is no invariant of the form $\bar{\xi}_\alpha C^{\alpha\beta} \eta_\beta$. This would imply $A^+ C A = C$ [instead of Eq. 1.46], which has no solution for C.

[2] For any 2×2 matrix M the identity

$$C^{-1} M^T C = M^{-1} \det M$$

can easily be verified, which for unimodular matrices reduces to Eq. 1.46.

[3] Equation 1.46 is the counterpart of Eq. 1.17. In the case of vectors, if $x'^\mu = L^\mu{}_\nu x^\nu$, then the covariant components x^μ transform according to L^{T-1}, or

$$x_\mu' = L_\mu{}^\nu x_\nu,$$

$$L_\mu{}^\nu = g_{\mu\sigma} L^\sigma{}_\varrho g^{\varrho\nu} \text{ (in matrix form } L^{T^{-1}} = GLG).$$

The matrices $A^{T^{-1}}$ form also a representation of the restricted Lorentz group; they are unimodular and the correspondence [1] $A \leftrightarrow A^{T^{-1}}$ preserves the multiplication law $(AB) \leftrightarrow (AB)^{T^{-1}} = A^{T^{-1}}B^{T^{-1}}$. The two representations A and $A^{T^{-1}}$ are equivalent[2] because

$$A = C^{-1}A^{T^{-1}}C. \tag{1.46'}$$

Thus covariant and contravariant spinors belong to two equivalent representation of L, $A^{T^{-1}}$ and A, respectively, both leaving the bilinear form Eq. 1.45 invariant.

There are two more two-dimensional representations that also leave the form Eq. 1.45 invariant, namely A^* and $A^{+^{-1}}$; i.e., altogether there are four matrices satisfying Eq. 1.46: A, $A^{T^{-1}}$, A^*, $A^{+^{-1}}$. In the case of 4 × 4 Lorentz transformations, the last two did not occur because L is real. Now we have two new representations which, however, are equivalent among themselves,

$$A^{+^{-1}} = CA^*C^{-1}, \tag{1.50}$$

which follows from Eq. 1.46, but these are *not* equivalent to the representation A. That is, there is no similarity transformation connecting A and A^*. (In mathematical language $A \to A^*$ is just an automorphism, *not* an inner automorphism, as was the case for $A \to A^{T^{-1}}$, or $A^* \to A^{+^{-1}}$.)

Therefore, we have new objects transforming according to the inequivalent representation A^*. It is convenient to denote these quantities by the same symbol but with dotted indices:

$$\xi'^{\dot\alpha} = A^{*\dot\alpha}{}_{\dot\beta}\,\xi^{\dot\beta}. \tag{1.51}$$

Now the covariant components transform according to $A^{+^{-1}}$. The metric tensor is $C^* = C$, and the invariant form

$$(\dot\xi, C\dot\eta) = \xi^{\dot\alpha}\eta_{\dot\alpha} = -\,\xi_{\dot\alpha}\eta^{\dot\alpha}. \tag{1.52}$$

Finally, let us look at the Lorentz transformations corresponding to these four different 2 × 2 matrices A, $A^{T^{-1}}$, A^*, $A^{+^{-1}}$. Let $L(A)$ be the transformation corresponding to $\pm A$, then, from Eq. 1.46,

$$L(A^{T^{-1}}) = L(C)L(A)L(C^{-1}).$$

---

[1] The mapping $A \leftrightarrow A^{T^{-1}}$, in mathematical language, is an inner automorphism.

[2] Two representations D and D' are called *equivalent* if there exists a fixed matrix B such that

$$D' = BDB^{-1};$$

in particular, we shall speak of unitary equivalence if B is unitary.

From Eq. 1.41 we find that $L(\pm C) = \sum_1 \sum_3$, where $\sum_j$ are the plane reflections perpendicular to the axes, Eq. 1.27. Similarly,

$$L(A^{+^{-1}}) = \sum_1 \sum_3 L(A^*) \sum_1 \sum_3 .$$

Because we do not have two inequivalent representations in the 4×4 L representation, we expect that $L(A^*)$ and $L(A^{+^{-1}})$ will be related to $L(A)$ and $L(A^{T^{-1}})$. Indeed, from Eq. 1.41 and using $\sigma_\mu^T = g_{\mu\nu}\sigma_2\tilde{\sigma}^\nu\sigma_2$, we get

$$L(A^*)^\mu{}_\nu = \tfrac{1}{2}\,\mathrm{tr}\,(\sigma_\mu A^* \sigma_\nu A^T) = \tfrac{1}{2}\,\mathrm{tr}\,[(A\sigma_\nu^T A^+ \sigma_\mu^T)^T] = \tfrac{1}{2}\,\mathrm{tr}(\sigma_\mu^T A \sigma_\nu^T A^+)$$

$$= \tfrac{1}{2}\,\mathrm{tr}[\tilde{\sigma}_\mu(\sigma_2 A \sigma_2)\tilde{\sigma}_\nu(\sigma_2 A\sigma_2)^+] ;$$

hence

$$L(A^*)^\mu{}_\nu = L(A^{T^{-1}})^\mu{}_\nu \det(\sigma_\mu\sigma_\nu)$$

and, similarly,

$$L(A^{+^{-1}})^\mu{}_\nu = L(A)^\mu{}_\nu \det(\sigma_\mu\sigma_\nu).$$

A representation D of the (Lorentz) group is called *reducible* if it has a semidiagonal or box form, or if it can be brought into this form by an equivalence transformation BDB^{-1}. In this case, the representation space is split into a number of lower-dimensional subspaces which are transformed into themselves under Lorentz transformation. Clearly it is sufficient to determine all *irreducible representations* of a group.

D. Higher-Order Spinor and Tensor Representations

From the two-dimensional representation $A(L)$ of the restricted Lorentz group, we can build up higher-dimensional representations successively. For example, in four dimensions we have quantities $\xi^{\alpha\beta}$, $\xi^{\dot{\alpha}\alpha}$, $\xi^{\dot{\alpha}\dot{\beta}}$ and the related quantities with covariant indices, which transform under Lorentz transformations as the product of two corresponding spinors, e.g.,[1]

$$\xi'^{\alpha\beta} = A^\alpha{}_\gamma A^\beta{}_\delta \xi^{\gamma\delta}, \qquad \xi'^{\dot{\alpha}\alpha} = A^{*\dot{\alpha}}{}_{\dot{\gamma}} A^\alpha{}_\gamma \xi^{\dot{\gamma}\gamma}, \qquad \xi'^{\dot{\alpha}\dot{\beta}} = A^{*\dot{\alpha}}{}_{\dot{\gamma}} A^{*\dot{\beta}}{}_{\dot{\delta}} \xi^{\dot{\gamma}\dot{\delta}}. \tag{1.53}$$

This process can be continued to any number of covariant and contravariant, dotted and undotted indices. A general spinor will transform as follows:

$$\xi'^{\alpha\beta\cdots\dot{\alpha}\dot{\beta}\cdots}_{ab\cdots\dot{a}\dot{b}\cdots} = A^\alpha_\sigma A^\beta_\rho \cdots A^{*\dot{\alpha}}_{\dot{\sigma}} A^{*\dot{\beta}}_{\dot{\rho}} \cdots A_a{}^s A_b{}^r \cdots A^*{}_{\dot{a}}{}^{\dot{s}} A^*{}_{\dot{b}}{}^{\dot{r}} \cdots \xi^{\sigma\rho\cdots\dot{\sigma}\dot{\rho}\cdots}_{sr\cdots\dot{s}\dot{r}\cdots}. \tag{1.54}$$

In this same manner we can generalize the vector transformations $x'^\mu = L^\mu{}_\nu x^\nu$ to arbitrary tensors; i.e., quantities $T^{\mu\nu}, \cdots$ that transform as the product of vectors. A general tensor transforms according to the law[1]

---

[1] $\xi^{\alpha\beta}$ transform according to the Kroenecker product $A \times A$, $\xi^{\dot{\alpha}\alpha}$ according to $A^* \times A$, etc.

$$T^{\mu'\nu\cdots}_{\sigma\rho\cdots} = L^{\mu}{}_{m}L^{\nu}{}_{n} \cdots L_{\sigma}{}^{s}L_{\rho}{}^{r} \cdots T^{mn\cdots}_{sr\cdots}. \tag{1.55}$$

The invariant form $\xi^{\alpha}C_{\alpha\beta}\xi^{\beta}$ generalizes to

$$\xi^{\alpha\cdots\dot{\alpha}\cdots}_{a\cdots\dot{a}\cdots} C_{\alpha\alpha'} \cdots C_{\dot{\alpha}\alpha} \cdots C^{aa'} \cdots C^{\dot{a}\dot{a}'} \cdots \xi^{\alpha'\cdots\dot{\alpha}'\cdots}_{a'\cdots\dot{a}'\cdots}, \tag{1.56}$$

As both the two-index spinors and the vectors refer to the four-dimensional representation of the Lorentz group, we must be able to relate the two. The general problem here is to construct tensors from the spinors that are more fundamental because of lower dimensionality.

The relation between vectors and two-index spinors is already apparent from the way we introduced the spinors in Section B. There we considered Hermitian 2 × 2 matrices of the form

$$X = \sigma_{\mu}x^{\mu}, \tag{1.57}$$

which transform under Lorentz transformations into

$$X' = AXA^{+}. \tag{1.58}$$

The matrix elements of X transform clearly like a two-index spinor, one dotted and one undotted; in fact Eq. 1.58 written in components reads

$$X'_{\alpha\dot{\alpha}} = A_{\alpha}{}^{\beta}A^{*\dot{\beta}}_{\dot{\alpha}}X_{\beta\dot{\beta}}, \qquad \text{or } A_{\alpha\beta}X^{\beta\dot{\beta}}A^{*}_{\dot{\alpha}\dot{\beta}}.$$

Therefore, Eq. 1.57 written in components connects vectors with second-rank mixed spinors:

$$X_{\alpha\dot{\alpha}} = \sigma_{\mu\alpha\dot{\alpha}}x^{\mu}. \tag{1.59}$$

$\sigma_{\mu\alpha\dot{\alpha}}$ are the matrix elements of the Pauli matrices (1.40) A Hermitian second-rank spinor is determined by four real numbers like the vector x^{μ}. Remembering that the spinor indices are raised and lowered by $C_{\alpha\beta}$, etc., we find the following relations connecting the metric tensor $g_{\mu\nu}$ and the spinor tensor $C_{\alpha\beta}$ and the mixed quantities $\sigma_{\mu\alpha\dot{\alpha}}$:

$$\begin{gathered} \sigma^{\alpha}_{\mu\gamma}\sigma^{\beta\dot{\gamma}}_{\gamma} + \sigma^{\alpha}_{\nu\dot{\gamma}}\sigma^{\beta\dot{\gamma}}_{\mu} = g_{\mu\nu}C^{\alpha\beta}, \\ \sigma^{\mu}_{\alpha\dot{\alpha}} = g^{\mu\nu}\sigma^{\gamma\dot{\delta}}_{\nu}C_{\gamma\alpha}C_{\dot{\delta}\dot{\alpha}}, \\ g_{\alpha\dot{\alpha}\beta\dot{\beta}} = C_{\alpha\beta}C_{\dot{\alpha}\dot{\beta}}. \end{gathered} \tag{1.60}$$

---

[1] Spinors and tensors may be defined under more general coordinate transformations. For example, under arbitrary coordinate transformations $x' = x'(x)$, tensors transform as

$$T^{\mu\nu}_{\sigma\varrho} ::: = \frac{\partial x'^{\mu}}{\partial x^{m}}\frac{\partial x'^{\nu}}{\partial x^{n}} \cdots \frac{\partial x^{s}}{\partial x'^{\sigma}}\frac{\partial x^{r}}{\partial x'^{\varrho}} \cdots T^{mn}{}_{sr} ::: . \tag{1.55'}$$

A general theory of tensor and spinor calculus is not required in this book, because we are dealing with Lorentz covariant theories only.

The general connection between the tensors and the spinors is then

$$X^{\gamma\dot\gamma}_{\alpha\dot\alpha\beta\dot\beta} = \sigma^{\lambda}_{\alpha\dot\alpha}\sigma^{\mu}_{\beta\dot\beta}X^{\ \nu}_{\lambda\mu}\sigma^{\gamma\dot\gamma}_{\nu}; \qquad X^{\nu}_{\lambda\mu} = \sigma^{\alpha\dot\alpha}_{\lambda}\sigma^{\beta\dot\beta}_{\mu}X^{\gamma\dot\gamma}_{\alpha\dot\alpha\beta\dot\beta}\sigma^{\nu}_{\gamma\dot\gamma}. \tag{1.61}$$

Thus spinor operations can be translated into the tensor operations and *vice versa*. Some examples of this translation will be discussed later under physical applications. For example, the quantity $\xi_\alpha\sigma^{\alpha\dot\alpha}_{\mu}\xi_{\dot\alpha}$, or in matrix notation $\xi\sigma_\mu\hat\xi$, will transform under proper Lorentz transformations as a vector. Instead of the dotted and undotted indices, we can work completely in matrix notation. Let us denote the spinors ξ and $\hat\xi$ transforming, respectively, as

$$\xi' = A\xi, \qquad \hat\xi' = A^*\hat\xi.$$

Let $\tilde\xi = C^{-1}\hat\xi$, then

$$\tilde\xi' = C^{-1}\hat\xi' = C^{-1}A^*\hat\xi = C^{-1}A^*C\tilde\xi = \tilde A\tilde\xi, \tag{1.62}$$

where $\tilde A = C^{-1}A^*C$.

When we write equations in matrix form they stand for both upper and lower indices; e.g., $\xi' = A\xi$ stands for $\xi^{\alpha\prime} = A^{\alpha}{}_{\beta}\xi^{\beta}$ and $\xi'_\alpha = A_\alpha{}^{\beta}\xi_\beta$. The elements A^{β}_{α} and A^{α}_{β} are different but related. Similarly, $\sigma_{\mu\alpha\dot\alpha}$ and $\sigma^{\dot\alpha}_{\mu\alpha}$, $\sigma^{\alpha\dot\alpha}_{\mu}$, $\sigma^{\alpha}_{\mu\dot\alpha}$ are different matrix elements. We choose $\sigma_{\mu\alpha\dot\alpha}$ to be the usual matrix elements.

Another way of introducing the higher-dimensional tensors and spinors is by writing down the general representations of the 2 × 2 unimodular group. Consider the spinors

$$\xi = \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}; \qquad \eta = \begin{pmatrix}\xi_{\dot 1}\\ \xi_{\dot 2}\end{pmatrix},$$

which transform with A and A^* respectively and form the normalized monomials[1]

$$M^{m_1m_2}_{j_1j_2} = \frac{\xi_1^{j_1-m_1}\xi_2^{j_1+m_1}\xi_{\dot 1}^{j_2-m_2}\xi_{\dot 2}^{j_2+m_2}}{[(j_1-m_1)!(j_1+m_1)!(j_2-m_2)!(j_2+m_2)!]^{1/2}}, \tag{1.63}$$

---

[1] This procedure is similar to that of finding the representations of the 2 × 2 unitary and unimodular group, i.e. the spinor corresponding to the three-dimensional rotation group. Explicitly, if the spinor $\xi = \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}$ transforms under the 2 × 2 unitary and unimodular group U, then the normalized monomials

$$M_j{}^m = \frac{\xi_1{}^{j-m}\xi_2{}^{j+m}}{[(j-m)!(j+m)!]^{1/2}}, \qquad -j < m < +j,$$

transform with $\mathscr{D}^{(j)}(U)$, the $(2j+1) \times (2j+1)$ matrix representation of U or R_3, the three-dimensional rotation group.

with $-j_1 < m_1 < +j_1$ and $-j_2 < m_2 < +j_2$ in integral steps. For given j_1 and j_2 the M's form components of a $(2j_1 + 1)(2j_2 + 1)$-dimensional vector. Under a Lorentz transformation these vectors transform according to

$$M' = \mathscr{D}^{(j_1 j_2)}(A)M, \tag{1.64}$$

where $\mathscr{D}^{(j_1 j_2)}$ is a $(2j_1 + 1)(2j_2 + 1) \times (2j_1 + 1)(2j_2 + 1)$ matrix and, in principle, can be worked out from Eq. 1.63. By varying j_1 and j_2 from 0 to ∞, we get all the representations of the unimodular group A. We have also

$$\mathscr{D}^{(j,0)} = \mathscr{D}^{(j)}.$$

(E) Infinitesimal Generators of the Lorentz and Unimodular Groups

We have seen that the inhomogeneous Lorentz transformations have ten parameters, the homogeneous ones six. These parameters occur in the matrix elements of L (six) and in the translation vector a (four), Eq. 1.20. There are other, sometimes more convenient, ways of exhibiting the parameters of the transformations. For example, we can write the proper Lorentz transformations in the form of an exponential matrix,[1]

$$L = \exp(a^{\mu\nu} M_{\mu\nu}), \tag{1.65}$$

where both $a^{\mu\nu}$ and $M_{\mu\nu}$ are antisymmetric in the indices ρ, σ, or in the form

$$L = \exp\left(\sum_{i=1}^{6} a_i K_i\right). \tag{1.66}$$

In both these expressions $a^{\mu\nu}$ and a_i are six real parameters and $I_{\mu\nu}$ and K_i are fixed 4×4 matrices.

More intuitively, we can use rotations around a space axis $\mathbf{n}$ by an angle θ

$$L = \exp(i\mathbf{J}\cdot\mathbf{n}\theta), \tag{1.67}$$

and pure Lorentz transformations by a hyperbolic angle χ [see Eq. 1.34],

$$L = \exp(i\mathrm{N}\cdot\mathbf{n}\chi), \tag{1.67$'$}$$

---

[1] The following properties of the exponential matrix may be readily verified:

(a) If K is a complex $n \times n$ matrix, the series $e^K = 1 + K + K^2/2! + K^3/3! + \cdots$ is always convergent;

(b) $\mathrm{Det}(e^K) = e^{\mathrm{tr}\, K}$;

(c) $e^{K+L} = e^K e^L$ if $[K, L] = 0$;

(d) $e^{SKS^{-1}} = Se^K S^{-1}$;

(e) If $\lambda_1, \cdots, \lambda_n$ are eigenvalues of K, then $e^{\lambda_1}, \cdots, e^{\lambda_n}$ are the eigenvalues of e^K;

(f) $e^{K*} = (e^K)^*$, $e^{K^T} = (e^K)^T$, $e^{K+} = (e^K)^+$, $e^{-K} = (e^K)^{-1}$;

(g) $e^K \leftrightarrow K$ is a mapping of a neighborhood of unity into a neighborhood of zero in the space of the parameters of K.

again exhibiting the six parameters and the six matrices **J** and **N**. In all these expressions the coefficients of the parameters are called the *infinitesimal generators* of the Lorentz group. If we expand the exponentials for infinitesimally small values of the parameters, and use the defining Eq. 1.17 of the Lorentz transformations, we get

$$K_i^+ G + G K_i = 0, \tag{1.68}$$

and similarly for the other form of the generators. The problem of the representations of the group can be reformulated in terms of the generators by finding all the matrices satisfying this last equation.

The different group generators do not commute, as the corresponding Lorentz transformations do not. These commutation relations can be worked out from the properties of the rotations. For example, for Eqs. 1.67 and 1.67′, one finds

$$\begin{aligned} [J^j, J^k] &= i\varepsilon_{jkl} J^l, \\ [J^j, N^k] &= i\varepsilon_{jkl} N^l, \\ [N^j, N^k] &= -i\varepsilon_{jkl} J^l, \end{aligned} \tag{1.69}$$

or, introducing a dual antisymmetric tensor by **J**: (M^{23}, M^{31}, M^{12}) and $\mathbf{N} = (M^{01}, M^{02}, M^{03})$, $M^{\mu\nu} = -M^{\nu\mu}$,

$$[M^{k\lambda}, M^{\mu\nu}] = -i(g^{\lambda\mu} M^{k\nu} + g^{k\nu} M^{\lambda\mu} - g^{k\mu} M^{\lambda\upsilon} - g^{\lambda\upsilon} M^{k\mu}). \tag{1.69′}$$

In this form we get back Eq. 1.65. $M^{\mu\nu}$ then correspond to rotations in the plane $(\mu\nu)$, with the matrix elements

$$(M^{\mu\nu})^{\sigma}{}_{\rho} = -i[g^{\mu\sigma} g^{\nu}_{\rho} - g^{\nu\sigma} g^{\mu}_{\rho}]. \tag{1.70}$$

For the inhomogeneous Lorentz group we have, in addition to L, the translation T, which can be written in the form

$$T = \exp(i a_\mu P^\mu), \tag{1.71}$$

where the generators P^μ commute among themselves, but not with $M^{\mu\nu}$,

$$[M^{\mu\nu}, P^\sigma] = -i(g^{\mu\sigma} P^\nu - g^{\nu\sigma} P^\mu). \tag{1.72}$$

From this one can evaluate $[\mathbf{J}, P^\mu]$ and $[\mathbf{N}, P^\mu]$.

Finally, we exhibit the generators of the 2 × 2 unimodular group. This is again a six-parameter group, and we can write

$$A = \exp[\tfrac{1}{4}(M^{\mu\nu} \sigma_\mu \sigma_\nu)] = \exp\left(\frac{i}{\mathbf{2}} \mathbf{j} \cdot \boldsymbol{\sigma} + \frac{1}{2} \boldsymbol{\lambda} \cdot \boldsymbol{\sigma}\right). \tag{1.73}$$

From Eq. (1.41), we see that $\exp(\tfrac{i}{2} \mathbf{J} \cdot \boldsymbol{\sigma})$ generates pure rotations and $\exp(\tfrac{1}{2} \boldsymbol{\lambda} \cdot \boldsymbol{\sigma})$ generates pure Lorentz transformations.

To summarize, an r-parameter group of $n \times n$ matrices will have r linearly independent generators, which form an r-dimensional vector space. The group property of the matrices implies that the commutator of any two generators $[K_1, K_2]$ also belongs to this vector space;[1] i.e., it is a linear combination of the other generators

$$[K_i, K_j] = \sum_l k_{ijl} K_l . \tag{1.74}$$

The coefficients k_{ijl} characterize the group and are called the *structure constants*. Equations 1.69 and 1.69′ are examples of this relation.

(F) Complex Lorentz Group

At the end of Section 2 we have given two generalizations of Lorentz space in complex domain. Correspondingly, two types of complex Lorentz groups may be defined by the equations

$$L^+ G L = G\,; \qquad L^T G L = G, \tag{1.75}$$

respectively. The first one leaves a real norm $(\bar{z}z)$ invariant, the second one a complex norm zz. We consider the second possibility.

In this section we shall write x, y for complex four-vectors. We consider again the 2×2 matrix

$$X = \sigma_\mu x^\mu,$$

which is no longer hermitian, as the x^μ's are complex, and the transformation

$$X' = AXB = \sigma_\mu x'^\mu, \qquad \det A = \det B = 1. \tag{1.76}$$

Hence

$$\det X' = \det X = x^\mu x_\mu = x^{\mu'} x_{\mu'}.$$

Thus the transformation, Eq. 1.76, generates a complex Lorentz transformation with $x^{0^2} - x^{1^2} - x^{2^2} - x^{3^2} =$ invariant, where x^μ are complex numbers. In the same manner as in Section (B), we find the Lorentz transformation $L(A, B)$ corresponding to two unimodular matrices A and B,

$$L(A, B)^\mu{}_\nu = \tfrac{1}{2} \operatorname{tr}(\tilde{\sigma}^\mu A \sigma_\nu B). \tag{1.77}$$

Equation 1.41 is a special case of this relation.[2]

---

[1] See exercise 7.

[2] For applications of the complex Lorentz group in quantum field theory, see R. Jost, *Helv. Phys. Acta*, *33*, (1960), 773; R. Jost, *Physics in 20th Century* (New York; Interscience Publishers, 1960). A. S. Wightman, *Dispersion Relations and Elementary Particles*, C. de Witt, ed. (New York; John Wiley and Sons, Inc., 1960).

Again the complex Lorentz group splits into two parts, one with det $= +1$, the other with det $= -1$. But now, in contrast to the real Lorentz group, $L_+^\uparrow$ and $L_+^\downarrow$ are continuously connected. For example, the complex Lorentz transformation

$$\begin{pmatrix} \cos\phi & i\sin\phi & 0 & 0 \\ i\sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix}, \qquad 0 \leqq \phi \leqq \pi,$$

connects continuously the identity transformation with L_{PT}, space and time reflection.

6. *The Principle of Relativity: Invariance and Covariance*

The inertial observers will compare their coordinates (the space-time events) with the help of Lorentz transformations. The observers must also compare the physical laws. The physical laws must be independent of the coordinate frames, i.e., the same for all inertial frames. In other words, it is impossible to infer a particular inertial frame from experiment, for example, whether or not a coordinate frame is at rest or in uniform motion. This is the *principle of relativity*. We shall use it in the form that all *physical laws must be invariant under the inhomogeneous Lorentz transformations.*

As a generalization of ordinary vector calculus (which expresses the fact that the vector equations are valid in any three-dimensional coordinate frame), we can write the fundamental equations of physical theories as four-vector, tensor or spinor equations such that the *form* of these equations is the same for all observers. In this form all terms of the equations transform in the same manner. We will then say that the equations are *covariant.*

However, it is important to realize that the *relativistic invariance* does not necessarily imply covariance. The equations need not be manifestly covariant; that is, the terms of the equations are not simply vectors, tensors or spinors; nevertheless, there is a definite law of transformation of these quantities. For example, one may use more general representations of the Lorentz group, then the tensor or spinor representations, those involving an additional transformation, if the theory in question will allow other invariance transformations besides the Lorentz invariance. (See Chapter V, Section 3.)

Or, an equation may not be manifestly covariant because it has been reduced to a particular frame. If a covariant equation reduces to a familiar equation in a particular frame, then the former is a unique covariant form of the latter, for if an equation is valid in one frame, it must be valid in all frames.

Examples of covariant and noncovariant, but invariant, equations will be given later in this book.

7. *Tensor and Spinor Fields and Momentum-Space Functions*

Particle properties are described by constant tensors or spinors in space, whereas fields are defined to be "objects" that are continuous functions of space–time or momenta of the particles. They can be classified according to their transformation properties:

(*a*) *Scalar fields*, $\phi(x)$. A scalar field describing some physical quantity has the same numerical value at each physical point P for all observers. If two observers denote the point P by x and x' and the quantity by ϕ and ϕ', we have

$$\phi'(x') = \phi(x). \tag{1.78}$$

In particular, there are *invariant scalar fields*; e.g., $f(x^2)$, $f(kx)$, whose functional form is also invariant under Lorentz transformations.

(*b*) *Vector field* $\phi_\mu(x)$. The physical vector quantity at some physical point P, again viewed by two different observers, will be denoted $\phi'_\mu(x')$ and $\phi_\mu(x)$, respectively. Although the physical vector is fixed, a component of this vector in one frame is a linear combination of the components with respect to the rotated frame under L. Hence

$$\phi^{\mu'}(x') = L^\mu{}_\nu \phi^\nu(x), \qquad \phi'_\mu(x') = L_\mu{}^\nu \phi_\nu(x), \tag{1.79}$$

for the covariant and contravariant components respectively.

Derivatives of a scalar field, $\partial\phi/\partial x^\mu \equiv \phi,_\mu$, transform like a vector field, for

$$\phi'_{,\mu}(x') = \frac{\partial\phi'(x')}{\partial x'^\mu} = \frac{\partial\phi(x)}{\partial x^\mu}\frac{\partial x^\nu}{\partial x'^\mu} = L_\mu{}^\nu \phi_{,\nu}(x)\,;$$

similarly, $\phi'^{,\mu}(x') = L^\mu{}_\nu \phi,^\nu(x)$.

(*c*) *Tensor field.* The physical quantity at each space-time point has an array of components. It is a tensor field if the transformation law is the generalization of Eq. 1.79:

$$T'^{\mu\cdots\sigma}_{\nu\cdots\rho}(x') = L^{\mu}{}_{m} \cdots L^{\sigma}{}_{s} L_{\nu}{}^{n} \cdots L_{\rho}{}^{r} T^{m\cdots s}_{n\cdots r}(x) \qquad \textbf{(1.80)}$$

Tensors of the same rank may be added or subtracted. If a covariant and a contravariant index of a tensor are equal, they may be contacted (i.e., these indices may be deleted; the resultant tensor is of rank $n - 2$. This is because the form $\phi^{\mu}\phi_{\mu}$ is an invariant scalar. The derivative of a vector field is a tensor field of rank two:

$$\phi'^{\mu}_{,\nu}(x') = L^{\mu}{}_{\sigma} L_{\nu}{}^{\rho} \phi^{\sigma}_{,\rho}(x),$$

which can be verified in the same manner as for $\phi_{,\mu}$.

In the same way we can verify that forms like

$$T^{\mu\cdots\sigma} a_{\mu} \cdots b_{\sigma}, \quad \text{or} \quad T^{\mu\cdots\sigma} S_{\mu\cdots\sigma},$$

are invariant, for

$$T'^{\mu\cdots\sigma}(x') S'_{\mu\cdots\sigma}(x') = \frac{\partial x'^{\mu}}{\partial x^{m}} \cdots \frac{\partial x'^{\sigma}}{\partial x^{s}} \frac{\partial x^{n}}{\partial x'^{\mu}} \cdots \frac{\partial x^{r}}{\partial x'^{\sigma}} T^{m\cdots s}(x) S_{n\cdots r}(x)$$

$$= \delta^{n}_{m} \cdots \delta^{r}_{s} T^{m\cdots s}(x) S_{n\cdots r}(x) = T^{m\cdots s}(x) S_{m\cdots s}(x).$$

In the case of Lorentz transformations, notice that we have set

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = L^{\mu}{}_{\nu}, \quad \frac{\partial x^{\nu}}{\partial x'^{\mu}} = (L^{-1})^{\nu}{}_{\mu} = (GL^{T}G)^{\nu}{}_{\mu} = g^{\nu\sigma} L_{\rho\sigma} g^{\rho}{}_{\mu} \equiv L_{\mu}{}^{\nu};$$

thus, in our notation,

$$L^{\mu}{}_{\sigma} L_{\mu}{}^{\rho} = \delta^{\rho}_{\sigma}.$$

(*d*) In terms of group representations we can express the general geometrical objects in the following form. Let $D(L)$ be an arbitrary representation of the Lorentz group; i.e., for two Lorentz transformations L_1 and L_2,

$$D(L_2 L_1) = D(L_2) D(L_1).$$

Let ξ^{α} be the vectors in the space in which $D(L)$ acts. Then we have fields transforming according to

$$\xi'^{\alpha}(x') = \xi'^{\alpha}(Lx) = D(L)^{\alpha}{}_{\beta} \xi^{\beta}(x).$$

Tensors are special examples of this transformation. Another class of quantity is the spinor field.

Spinor field. In matrix form the spinor field transforms as

$$\xi'(x') = S(L)\xi(x) \tag{1.81}$$

where $S(L)$ is a general spin representation of the Lorentz group.

The previous transformation properties also apply to functions of other variables, for example, functions of the momenta of the particles. We have again tensors

$$T^{\mu\cdots\sigma}_{\nu\cdots\rho}(k),$$

and spinors

$$\xi^{\alpha}(k),$$

transforming according to Eqs. 1.80 and 1.81 respectively, which are functions of one or more four-vectors k.

8. *Analysis*

In this section we shall discuss some of the differentiation and integration properties of fields and surfaces in four-dimensional space.

We introduce the four-dimensional gradient and divergence operations by

$$\partial\phi \equiv \partial_\mu \phi \equiv \phi_{,\mu} = \frac{\partial\phi(x)}{\partial x^\mu},$$

$$\partial_\mu \phi^\mu = \partial^\mu \phi_\mu = \frac{\partial\phi^\mu}{\partial x^\mu} \equiv \phi^\mu_{,\mu};$$

similar expressions hold for other covariant or contravariant indices. Note that $\partial_\mu = \partial/\partial x^\mu$ is a covariant and $\partial^\mu = \partial/\partial x_\mu$ is a contravariant vector.

The gradient vector of a scalar field $\partial_\mu \phi$ is "orthogonal" to the three-dimensional surface in the sense of our metric, for

$$\partial_\mu \phi \, dx^\mu = d\phi = 0;$$

if dx^μ lies on the surface, ϕ is a constant.

The covariant generalization of the "curl" of a vector field is the antisymmetric tensor

$$G_{\mu\nu}(x) = \partial_\mu \phi_\nu(x) - \partial_\nu \phi_\mu(x) = -G_{\nu\mu}(x).$$

In three dimensions this reduces to the usual definition of the curl, $G_{ij} = \partial_i \phi_j - \partial_j \phi_i$, with $G_{23} = (\nabla \times \phi)_1$, plus two other equations obtained by cyclical permutations.

From the vector ∂_μ we can form formally and operationally a scalar, its "norm,"

$$\partial^\mu \partial_\mu \equiv \partial^2 \equiv \Box^2 = \frac{\partial^2}{\partial x^{0 2}} - \Delta^2 ;$$

$\Box^2$ is an invariant differential operator.

The curl of a gradient is always zero, for

$$\partial_\mu \phi_{,\nu} - \partial_\nu \phi_{,\mu} = 0 ;$$

however, the divergence of the curl is not zero. It will be written as

$$G^{;\nu}_{\mu\nu} \quad \text{or} \quad G^{\mu\nu}_{,\nu} = \partial^\mu \phi^\nu_{,\nu} - \Box^2 \phi^\mu .$$

The integral theorems of vector analysis can be generalized to four dimensions also.

The vanishing of the closed line integral $\oint \phi_\mu dx^\mu = 0$ implies $\phi_\mu = \phi_{,\mu}$, for $\oint \phi_{,\mu}\, dx^\mu = \oint d\phi = 0$.

The four-dimensional volume element will be written as

$$dx = dx^0\, dx^1\, dx^2\, dx^3 .$$

The boundary B of a four-dimensional region R is a three-dimensional surface which can be described either by three parameters u_1, u_2, u_3, or by its equations, $F(x^0, x^1, x^2, x^3) = 0$. The surface element on B can be represented by a four-vector $d\sigma^\mu$ defined by

$$d\sigma_\mu : (dx^1\, dx^2\, dx^3,\, dx^0\, dx^2\, dx^3,\, dx^0\, dx^1\, dx^3,\, dx^0\, dx^1\, dx^2).$$

If B is the ordinary three-dimensional space for $dx^0 = 0$, or $x^0 = \text{const}$, then $d\sigma_\mu$ is in the direction of $n_\mu : (1, 0, 0, 0)$. If we introduce the three parameters u_1, u_2, u_3 on B, Gauss' theorem reads

$$\int_R \frac{\partial \phi^\mu}{\partial x^\mu}\, dx = \int_B \begin{vmatrix} \phi^0 & \phi^1 & \phi^2 & \phi^3 \\ \dfrac{\partial x^0}{\partial u^1} & \dfrac{\partial x^1}{\partial u^1} & \dfrac{\partial x^2}{\partial u^1} & \dfrac{\partial x^3}{\partial u^1} \\ \dfrac{\partial x^0}{\partial u^2} & \dfrac{\partial x^1}{\partial u^2} & \dfrac{\partial x^2}{\partial u^2} & \dfrac{\partial x^3}{\partial u^2} \\ \dfrac{\partial x^0}{\partial u^3} & \dfrac{\partial x^1}{\partial u^3} & \dfrac{\partial x^2}{\partial u^3} & \dfrac{\partial x^3}{\partial u^3} \end{vmatrix} du^1\, du^2\, du^3 = \int_B d\sigma_\mu \phi^\mu, \quad \textbf{(1.82)}$$

where we used the relations

$$d\sigma_0 = \frac{\partial(x^1 x^2 x^3)}{\partial(u^1 u^2 u^3)}\, du^1\, du^2\, du^3 = dx^1\, dx^2\, dx^3, \text{ etc.}$$

Thus $d\sigma_\mu$ are independent of the parametrization of the surface B.

Of particular importance will be the notion of a *space-like surface.* B is space-like if $d\sigma^\mu$ is everywhere a time-like vector. A space-like surface can be defined also as the set of all vectors $x, y, \cdots$, whose "distance" $(x-y)^2$ is space-like. As we know, the normal vector to a three-dimensional space-like manifold is time-like. The only difference between three-dimensional space and a space-like surface is that in the latter case the normal four-vector to the surface may be an arbitrary time-direction which is equivalent to the x^0 direction. Points of space-like surface cannot be reached from each other by any signal. A space-like surface passing through the origin will thus consist of vectors y which are orthogonal to a time-like vector n: $ny = 0$. An arbitrary vector x has a "component" on this surface given by $y = x - n(nx)$. [See Eqs. 1.9 and 1.9′.] We will also introduce space-like surfaces not passing through the origin. These will be characterized by vectors y satisfying $ny = \theta$, where θ is a real parameter. For, if two vectors y_1 and y_2 satisfy $ny_1 = ny_2 = \theta$, then $n(y_1 - y_2) = 0$, or all the differences $(y_1 - y_2)$ are orthogonal to n. The spatial component of an arbitrary vector x is now

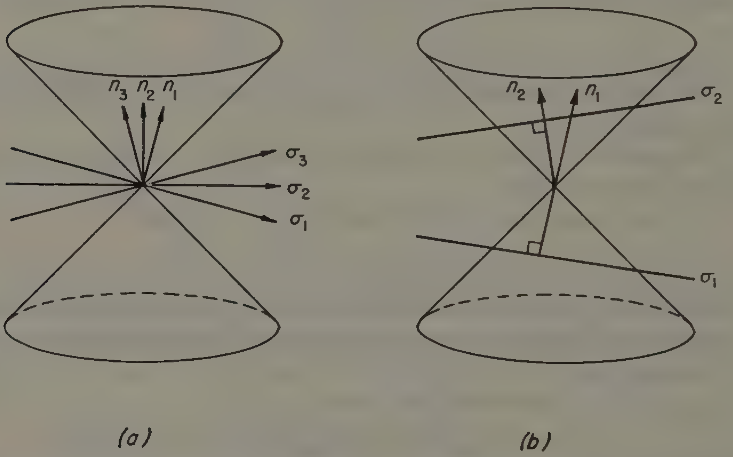


Fig. 2. (a) Space-like surfaces passing through the origin and their "normals"; (b) Space-like surfaces not passing through the origin

$$y = x - n(nx - \theta).$$

θ will play the role of a coordinate invariant time variable. (Fig. 2.)

From Gauss' theorem it follows that for a divergence-free vector, i.e., $\phi^\mu_{;\mu} = 0$, we have the relation

$$\int_B d\sigma_\mu \phi^\mu = 0.$$

If we now take B to consist of any two space-like surfaces σ_1 and σ_2, and if we assume that ϕ^μ approaches zero sufficiently rapidly, then we can neglect the contribution to B of the infinite cylinder connecting σ_1 and σ_2, and we get

$$0 = \int_B d\sigma_\mu \phi^\mu = \int_{\sigma_2} d\sigma_\mu \phi^\mu - \int_{\sigma_1} d\sigma_\mu \phi^\mu,$$

or, $\int_\sigma d\sigma_\mu \phi^\mu$ is independent of the space-like surface σ, not only at a given time, but at all times. In other words, under the stated conditions, $\int d\sigma_\mu \phi^\mu$ is a constant, for which we can also write, choosing, say, the surface $t = 0$,

$$\int d\sigma_\mu \phi^\mu = \int \phi^0 \, d^3x = \text{const.} \tag{1.83}$$

The converse of this theorem is also true: if $\int_\sigma d\sigma_\mu \phi^\mu$ is independent of σ, then $\phi^\mu_{;\mu} = 0$. $\int d\sigma_\mu \phi^\mu$ is independent of σ, then $\phi^\mu_{;\mu} = 0$. $\int d\sigma_\mu \phi^\mu$ is then a scalar number.

9. *Further Developments and Exercises*

1. Consider all linear, nonsingular transformations that leave $x^{0^2} - r^2 = 0$ invariant and that map origin into origin. If we do not invoke the dimensional argument, what other restrictions are necessary to derive Eq. 1.3 in text? (See V. Fock, *The theory of space-time and gravitation*, Pergamon Press, 1959.)

2. Discuss the eigenvalues of complex Lorentz transformation. [See E. P. Wigner, *Annals of Mathematics*, *40*, (1939), 149.]

3. Verify Eqs. (1.31) and (1.33).

4. *Appearance of moving objects.* The eye or the camera looking at a moving object records photons arriving at the same time. Due to the motion of the object, these photons cannot have been emitted at the same time. Show that for an object subtending a small solid angle this distortion of the object cancels, to a first approximation, the distortion due to Fitzgerald-Lorentz contraction so that, say, a moving circular object will appear circular again.

5. *The clock "paradox"*. Innumerable papers have been published on this famous problem.[1] To avoid the complication due to the turning of

---

[1] For some recent expositions, see A. Schild, *Amer. Mathem. Monthly*, *66*, (1959), 1; G. A. Blass, *Amer. Mathem. Monthly*, *67*, (1960), 754.

velocity, we modify the problem as follows: Consider three points on a straight line A, B, C, the distance AB being equal to the distance BC. A light signal is sent from B. At the arrival of the signal at A and C, two travelers move from A and C towards B with the same speed. If, in the fixed coordinate frame at A, the time needed to travel from A to B is t, the time needed in the moving frame is $t' = t(1-\beta^2)^{1/2}$. However, if we consider the moving frame as the primary one, the point A moves backward with the velocity $-v$. The distance AB will appear Lorentz-contracted to a traveler in a space ship, $x' = x(1-\beta^2)^{1/2}$. Again measured from the space ship, $x' = vt'$, or $t' = (x/v)(1-\beta^2)^{1/2}$. Therefore, with respect to both coordinate frames, the time t' is shorter, and there is no paradox. When the traveler reaches B, we concentrate on the traveler coming from C; the "return" trip is symmetric. If we consider a round-trip instead of the point C, the argument is the same. The effects of acceleration at the turning points can be neglected, the problem is purely a special relativistic one.

6. Prove that the determinant of $L(A)$ and $L(A, B)$ defined by Eqs. 1.41 and 1.77 is equal to $+1$. *Hint*: Start from the equation [see before Eq. 1.41],

$$L^{\mu}{}_{\nu}\sigma_{\mu\alpha\dot\alpha} = A_{\alpha}{}^{\beta}\sigma_{\nu\beta\dot\beta}A^{*\dot\beta}_{\dot\alpha}),$$

which can be written as $(A \times A^*)T = TL$, where

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} \sigma_{0,11} & \sigma_{1,11} & \sigma_{2,11} & \sigma_{3,11} \\ \sigma_{0,12} & \sigma_{1,12} & \sigma_{2,12} & \sigma_{3,12} \\ \sigma_{0,21} & \sigma_{1,21} & \sigma_{2,21} & \sigma_{3,21} \\ \sigma_{0,22} & \sigma_{1,22} & \sigma_{2,22} & \sigma_{3,22} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}.$$

Hence

$$A \times A^* = TLT^{-1},$$

and .

$$|\mathrm{tr}\, A|^2 = \mathrm{tr}\, L,$$

$$\det(A \times A^*) = (\det A)^4 = \det L.$$

[Note that $\det(A \times B) = (\det A)^{\dim B}(\det B)^{\dim A}$.]

7. Obtain the inversion of Eq. 1.41,

$$\pm A = \frac{L_{\mu\nu}\sigma^{\mu}\tilde\sigma^{\nu}}{(L_{\alpha\beta}L_{\gamma\delta}\sigma^{\alpha}\tilde\sigma^{\beta}\sigma^{\delta}\tilde\sigma^{\gamma})^{1/2}} = \frac{1}{N}\{\mathrm{tr}\, L + (L^{j}{}_{0} + L^{0}{}_{j} - i\varepsilon^{ijk}L^{k}{}_{l})\sigma_{j}\},$$

where

$$N^2 = 4 + \mathrm{tr}(LL) - (\mathrm{tr}\, L)^2 + i\varepsilon^{\mu\nu\sigma\lambda}L_{\mu\nu}L_{\sigma\lambda}.$$

8. We have introduced the matrices, $\tilde{\sigma}_\mu:(1, -\boldsymbol{\sigma})$. Prove the following relations:

$$\sigma^\mu\tilde{\sigma}^\nu = g^{\mu\nu} + \tfrac{1}{2}i\varepsilon^{\mu\nu\lambda\rho}\sigma_\lambda\tilde{\sigma}_\rho ;$$

$$\tilde{\sigma}^\mu\sigma^\nu = g^{\mu\nu} - \tfrac{1}{2}i\varepsilon^{\mu\nu\lambda\rho}\tilde{\sigma}_\lambda\sigma_\rho ;$$

$$[\sigma^\mu, \tilde{\sigma}^\nu] = \tfrac{1}{2}\, i\varepsilon^{\mu\nu\lambda\rho}[\sigma_\lambda\tilde{\sigma}_\rho - \tilde{\sigma}_\lambda\sigma_\rho]; \qquad \{\sigma^\mu, \tilde{\sigma}^\nu\} = 2g^{\mu\nu} \quad + \tfrac{1}{2}\varepsilon^{\mu\nu\lambda\rho}[\sigma_\lambda\tilde{\sigma}_\rho + \tilde{\sigma}_\lambda\sigma_\rho]$$

$$\sigma^\mu\tilde{\sigma}^\nu + \sigma^\nu\tilde{\sigma}^\mu = 2g^{\mu\nu}$$

$$\mathrm{tr}[\sigma^\mu\check{\sigma}^\nu] = 2g^{\mu\nu} ;$$

$$\mathrm{tr}[\sigma^\mu\tilde{\sigma}^\nu\sigma^\lambda\tilde{\sigma}^\rho] = 2[g^{\mu\nu}g^{\lambda\rho} - g^{\mu\lambda}g^{\nu\rho} + g^{\mu\rho}g^{\nu\lambda} - i\varepsilon^{\mu\nu\lambda\rho}] ;$$

$$\sigma^\mu_{\alpha\dot{\alpha}}\tilde{\sigma}_{\mu\beta\dot{\beta}} = 2\delta_{\alpha\beta}\delta_{\dot{\alpha}\dot{\beta}} .$$

9. *Covering group.* Show that the group $SL(2, c)$ is the universal covering group of the proper Lorentz group $L_+^\uparrow$. (See Wightman, *loc. cit.*)

10. Prove Eq. 1.74.

11. Show that the three-dimensional complex orthogonal group 0_3 is isomorphic to the proper Lorentz group $L_+^\uparrow$ and that the corresponding formulas (as in Exercises 6 and 7) are

$$R_{ij} = \tfrac{1}{2}\,\mathrm{tr}[\sigma_i A \sigma_j A^{-1}] ;$$

$$\pm A = \frac{1}{N}(1 + \sigma_i\sigma_j R_{ij}) ; \qquad N^2 = 4(1 + \mathrm{tr}\, R).$$

12. Obtain the connection between the parameters of the Lorentz group in $x'_\mu = L_\mu{}^\nu x_\nu + a_\mu$ and that associated with the infinitesimal generators

$$x' = e^{\zeta_{\mu\nu}X^{\mu\nu} + \pi_\mu p^\mu} x.$$

13. Discuss Lorentz transformations when curvilinear coördinates are used.

14. Consider the four-vector

$$W^\lambda = \tfrac{1}{2}\varepsilon^{\lambda\mu\nu\rho}p_\mu J_{\nu\rho} ;$$

Show that W^2, like $p^2 = p_\mu p^\mu$, commute with all the infinitesimal operators. Show that

$$[p_\lambda, W_\mu] = 0, \qquad [W^\mu, W^\nu] = -i\varepsilon^{\mu\nu\sigma\rho}p_\sigma W_\rho .$$

For a given momentum p, $p^2 = m^2$, $m \neq 0$, choose a basis, as in Eq. 1.13, with vectors

$$e_0 = \mathbf{p}/m,\ e_1,\ e_2,\ e_3,$$

satisfying Eq. 1.14. Let

$$W^{\lambda} = \sum_{i=1}^{3} S^i e_i^{\lambda},$$

because $W^{\lambda} p_{\lambda} = 0$; then

$$W^2 = W_{\mu} W^{\mu} = -\sum_{T}^{3} S^{i^2},$$

and the three operators S^i have the commutation relations of the angular momentum

$$[S^i, S^j] = im\varepsilon^{ijk} S^k.$$

This provides a covariant description of spin or polarization for particles of nonzero mass. The S^i are the infinitesimal generators of the three-dimensional orthogonal group which has finite-dimensional representations such that

$$W^2 = -S^2 = -j(j+1)m,$$

and $\mathbf{W} \cdot \boldsymbol{\eta}$ has eigenvalues m, $-j < m < +j$; j is an integral or half-odd integral number.

15. Show that unimodular matrix A corresponding to a Lorentz transformation, which takes a vector k into the direction $(m, 0, 0, 0)$ (i.e., rest frame), is given by

$$A^{-1} = \left(\frac{k^{\mu}\sigma_{\mu}}{m}\right)^{1/2}, \qquad A = \left(\frac{k^{\mu}\tilde{\sigma}_{\mu}}{m}\right)^{1/2},$$

the Hermitian square root of $(k^{\mu}\sigma^{\mu})/m$.

Show that for a four-vector, with $k^2 = m^2$,

$$(k^{\mu}\sigma_{\mu}/m)^{1/2} = \cosh(\chi/2) + \hat{\mathbf{k}} \cdot \boldsymbol{\sigma} \sinh(\chi/2),$$

where $\hat{\mathbf{k}}$ is the unit three-vector in the spatial direction and $\cosh \chi = k^0/m$.

16. Solve the same problem for light-like and space-like vectors when they are transformed into the vectors $(m, 0, 0, m)$ and $(0, 0, 0, m)$, respectively.

17. Determine the group of Lorentz transformations that transform a given four-vector into itself. (*The little group of the four-vector.*) Show that for a time-like vector the A's of the little group are unitary and that the little group is isomorphic to the three-dimension rotation group.

Show that the little group of a light-like vector is essentially isomorphic to the group of rotations in a Euclidian plane.

18. Show that any non-singular matrix A can be decomposed in one and only one way in the form $A = HU$ where H is positive definite hermitian and U is unitary. Apply this theorem to 2×2 unimodular matrices and show that U generates a rotation and H a pure Lorentz transformation. Give their explicit forms and prove the decomposition, Eq. 1.32′.

BIBLIOGRAPHY FOR CHAPTER I

Books

1. J. Aharoni, *The Special Theory of Relativity* (Oxford: Clarendon Press, 1959).
2. P. G. Bergmann, "The Special Theory of Relativity," *Handbuch der Physik* (Berlin: Springer, 1962), Vol. IV.
3. E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave Equations* (London: Blackie, 1953).
4. V. Fock, *The Theory of Space-Time and Gravitation* (London: Pergamon Press, 1959).
5. R. Jost, "Das Pauli-Prinzip und die Lorentz-Gruppe," *Theoretical physics in the twentieth century*, W. Pauli Memorial Volume, M. Fierz and V. Weisskopf, eds. (New York: Interscience, 1960).
6. A. Lichnerowitz, *Théories Relativistes de la Gravitation et de l'Électromagnetisme* (Paris: Mason & Cie., 1955).
7. C. Møller, *The Theory of Relativity* (London: Oxford University Press, 1952).
8. W. Pauli, *Theory of Relativity* (New York: Pergamon Press, 1958).
9. A. Papapetrou, *Spezielle Relativitätstheorie* (Berlin: Deutscher Verlag der Wissenschaften, 1957).
10. J. L. Synge, *Relativity: the Special Theory* (Amsterdam: North Holland Publishing Co., 1956).
11. J. L. Synge, *Relativity: the General Theory* (*Ibid.*, 1960).
12. A S. Wightman, in *Dispersion Relations and Elementary Particles*, C. de Witt and R. Omnés, eds. (New York: John Wiley & Sons, 1960).

Others

W. I. Bade and H. Jehle, "An Introduction to Spinors," *Rev. Mod. Phys.*, *25*, (1953), 714.

A. J. Macfarlane, "On the restricted Lorentz group and groups homomorphically related to it," *J. Math. Phys.*, *3*, (1962), 1116.

A. Schild, "Discrete Space-Time and Integral Lorentz Transformations," *Can. J. Math.*, *1*, (1949), 29.

L. I. Foldy, "Synthesis of Covariant Particle Equations," *Phys. Rev.*, *102*, (1956), 568.

II.

RELATIVISTIC DYNAMICS

1. *Proper Time Form of the Equations of Motion*

We start with the notion of *the mass point* and shall find its laws of motion in accordance with the principle of relativity and with Lorentz transformations. As we have discussed at the beginning of Chapter I, we expect to find a generalization of the equations of the classical mechanics that will reduce to the classical case in the limit of small velocities.

We cannot consider the time t as a parameter in the same way as in Newtonian mechanics, for Lorentz transformations involve also a change of t. However, this does not mean that we cannot differentiate with respect to t. In each frame one can still define velocity and acceleration and so on, as we shall see. A formally simple way, which parallels exactly the development of classical mechanics, is the introduction of an *invariant parameter s* which will play the role of t. We discuss this formalism first.

The trajectory of the mass point in space-time will be described by the "position" vector $x^\mu(s)$ as a function of the invariant parameter s to be specified presently. The trajectory in the four-dimensional Lorentz space will also be referred to as the *world line* of the mass point. As the invariant parameter, we first choose s defined by

$$ds^2 = dx^\mu \, dx_\mu = dx^{0^2} - dx^{1^2} - dx^{2^2} - dx^{3^2}; \tag{2.1}$$

later we shall also use arbitrary parameters τ. ds^2 is the simplest invariant one can form, and it has a definite physical meaning: In a coordinate system in which the particle is at rest $d\mathbf{x} = 0$, and ds^2 reduces to

$$ds^2 = dx^{0^2} = c^2 dt^2.$$

Thus, in the rest frame of the particle s is essentially the time, apart from the constant factor c^2. We shall denote s the *proper time* of the particle. The world line of the mass point in the rest frame is just the x^0 axis. In any other frame the value of s, not the meaning, will be the same as in the rest frame.

Now we can define velocities and accelerations with respect to the invariant parameter. The *four-velocity* is defined by

$$u = \frac{dx(s)}{ds}, \text{ or, in components, } u^\mu = \frac{dx^\mu(s)}{ds} = \dot{x}^\mu(s). \tag{2.2}$$

The four-velocity u is always a unit time-like vector, for

$$uu = \frac{dx}{ds}\frac{dx}{ds} = \frac{dx^2}{ds^2} = 1. \tag{2.3}$$

The components of u can also be written as

$$u^0 = \frac{dx^0}{ds}, \qquad \mathbf{u} = \frac{dx^0}{ds}\frac{1}{c}\mathbf{v},$$

where $\mathbf{v} = d\mathbf{x}/dt$ is the ordinary velocity. The three space components of $\mathbf{u}$ are the velocities in the frame of reference considered, the time component is something new. This can be obtained from Eq. 2.1:

$$u^0 = \frac{dx^0}{ds} = \frac{1}{(1 - v^2/c^2)^{1/2}} = \frac{1}{(1 - \beta^2)^{1/2}} = \gamma, \quad \text{or} \quad \frac{ds}{dt} = \frac{c}{\gamma}. \tag{2.4}$$

Similarly, we define an *acceleration* by

$$\ddot{x} = \dot{u} = \frac{du}{ds}. \tag{2.5}$$

The four-vector $\ddot{x}$ is always orthogonal to the velocity $\dot{x} = u$. For, differentiating Eq. 2.3, $u^2 = 1$, with respect to s we get

$$\dot{u}u = 0. \tag{2.6}$$

Thus we have in the 4-dimensional Lorentz space an analogue of the circular motion where velocity and acceleration are always perpendicular to each other. Both are consequences of the condition $x^2 = \text{const}$.

Finally, we introduce the *relativistic momentum*. Here we need a new parameter characterizing the mass point, *the rest mass* m_0. It is an intrinsic and invariant constant influencing the motion of the mass point. Its importance lies in the fact that different particles can be approximated by a mass point, as far as their motion is concerned, with different values of m_0. Otherwise, for the concept of abstract mass point, m_0 can be put equal to one. From experience it has been found sufficient to specify different particles by a single parameter within some limits. We shall see later that other intrinsic parameters, such as electric and magnetic moments and spin, must be introduced also. The explanation of the nature of these parameters, such as the rest mass m_0, is outside the scope of mechanics, which concerns itself with the laws of motion. Their explanation must be found in the theory of the structure of matter or of particles.

The four-momentum is defined by

$$p = m_0 c u, \tag{2.7}$$

or, in components,

$$p^0 = \frac{m_0 c}{(1-\beta^2)^{1/2}}, \qquad \mathbf{p} = \frac{m_0 \mathbf{v}}{(1-\beta^2)^{1/2}}. \tag{2.8}$$

The momentum p is always a time-like vector with norm

$$p^2 = m_0^2 c^2 = p^{0^2} - \mathbf{p}^2, \tag{2.9}$$

which follows from Eq. 2.3. Putting, formally first, $m = m_0/(1-\beta^2)^{1/2}$, the components of p take the form $p^0 = mc$ and $\mathbf{p} = m \cdot \mathbf{v}$; i.e., the usual classical space components of the momentum. The physical significance of this relation will be discussed below. Note that in the limit $v/c \to 0$ we get $\mathbf{p} = m_0 \mathbf{v}$.

Having defined the necessary concepts of the motion, we proceed now to the formulation of the law of motion. We see from Eq. 2.9 that, given m_0 (positive), the motion must be such that the vector p will always lie in

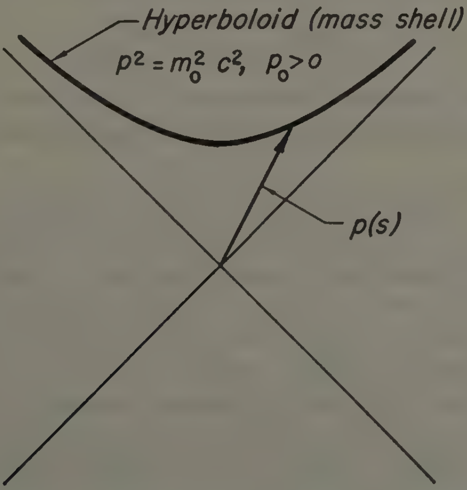


Fig. 3. The most general motion of a particle is described by a function $p(s)$ on the forward mass shell hyperboloid. p is the momentum, s the proper time

the forward hyperboloid shown schematically in Fig. 3. Because Lorentz transformations also keep the vector $p(s)$ on the hyperboloid, special types of (linear) "motions" will be obtained by the Lorentz transformations L, which take the vector $p(s)$ into some other vector $p(s')$. But we want to discuss the motion in a physical sense such that in the limit we get Newton's equations.[1]

The simplest *invariant* generalization of Newton's equations is

$$\dot{p} = \frac{dp}{ds} = K, \tag{2.10}$$

where K is the *four-force*. We see from Eq. 2.6 that K is always orthogonal to the velocity,

$$Ku = 0. \tag{2.11}$$

We shall refer to Eq. 2.10 as the *Minkowski equation* of motion.

We have not yet specified the form of K which characterizes the rate of change of the momentum. However, let us first see the physical significance of Eqs. 2.10 in the more familiar form by showing that these equations contain, in the limit, the three Newton's equations plus the energy equation.

Using Eq. 2.7, we write Eq. 2.10, viz.,

$$\frac{d}{ds}\left(m_0 c \frac{dx^\mu}{ds}\right) = K^\mu,$$

in the following form:

$$\frac{d}{dt}\left(m_0 c \frac{dt}{ds}\frac{dx^\mu}{dt}\right) = K^\mu \frac{ds}{dt}.$$

Comparing the three space components ($\mu = 1, 2, 3$) of this equation with the classical equation $(d/dt)(m\, d\mathbf{x}/dt) = \mathbf{F}$, we get a formal equality if we put

$$m_0 c \frac{dt}{ds} = \frac{m_0}{(1-\beta^2)^{1/2}} \overset{\text{def}}{\equiv} m, \quad \text{and} \quad \mathbf{K}\frac{ds}{dt} = \mathbf{F}. \tag{2.12}$$

There is more than a formal analogy. In the limit $v/c \to 0$, both equations become identical, and we see how to obtain the relativistic form of the Newtonian forces. Eq. 2.12 has also been interpreted by saying that the "mass" increases with velocity. But there is no internal mechanism increasing the mass of the particle; the really physical mass m_0 does not change. What Eq. 2.12 says is that we could correct the classical equation by making the mass variable according to Eq. 2.12. But Eq. 2.10 is the more fundamental formulation of the relativistic generalization.

---

[1] Even classical mechanics can be put into a form of transformation theory.

Let us return to the time component of Eq. 2.10:

$$\frac{d}{ds}\left(m_0 c\,\frac{dx^0}{ds}\right) = \frac{d}{ds}\,\frac{m_0 c}{(1-\beta^2)^{1/2}} = K^0.$$

From Eq. 2.11, we find K^0: $K^0\,dx^0/ds = \mathbf{K}\cdot\mathbf{u}$. Or, using Eq. 2.12,

$$\frac{d}{dt}(mc^2) = \mathbf{F}\cdot\mathbf{v}.$$

We can integrate this equation, obtaining

$$mc^2 = \int \mathbf{F}\cdot\mathbf{v}\,dt = \int \mathbf{F}\cdot d\mathbf{r}.$$

If we take conservative forces; i.e., $\mathbf{F} = -\nabla\phi$, we find[1]

$$mc^2 + \phi = \text{const} = E. \qquad \textbf{(2.13)}$$

We shall interpret the constant E as the total relativistic energy, thereby obtaining also the conservation law for the total energy. This is based on the limiting form of Eq. 2.13 as $v/c \to 0$. Expanding Eq. 2.13 in powers of v/c, we get

$$E = m_0c^2 + \tfrac{1}{2}m_0v^2 + \tfrac{3}{8}m_0\,\frac{v^4}{c^2} + \cdots + \phi. \qquad \textbf{(2.13')}$$

The total energy must be defined with the additive constant m_0c^2, called the *rest energy* in order to satisfy the conservation law, Eq. 2.13. The nonrelativistic definition of the energy is obtained as follows:

$$E_{NR} = \lim_{v/c\to 0}(E - m_0c^2). \qquad \textbf{(2.14)}$$

The quantity $T = (E - m_0c^2 - \phi)$ can be defined as the *relativistic kinetic energy.*[2]

---

[1] Or, we can integrate directly the equation $(d/dx)(mc^2) = cK^0 = (ds/dt)\mathbf{K}\cdot\mathbf{u}$ with respect to s;

$$mc^2 = \int \frac{ds}{dt}\,\mathbf{K}\cdot\mathbf{u}\,ds = \int \frac{ds}{dt}\,\mathbf{K}\cdot d\mathbf{r} = \int -\nabla\phi^L\cdot d\mathbf{r} = -\phi + E,$$

by Eq. 2.12.

[2] The expression for T can be obtained also from the following argument: $\mathbf{F}\cdot d\mathbf{r}$ is the work done by the external forces on the particle and is equal to the increase of its kinetic energy;

$$\mathbf{F}\cdot d\mathbf{r} = \frac{d}{dt}(m\mathbf{v})\,d\mathbf{r} = dT,$$

or,

$$dT = v^2\,dm + m\mathbf{v}\cdot d\mathbf{v} = \frac{m_0\mathbf{v}\cdot d\mathbf{v}}{(1-\beta^2)^{1/2}},$$

where we have obtained dm from Eq. 2.12. Integrating this equation with respect to $\mathbf{v}$ between 0 and v, we get

$$T = m_0c^2\left[\frac{1}{(1-\beta^2)^{1/2}} - 1\right],$$

which is just $(E - m_0c^2 - \phi)$.

Now we shall express E as a function of the momentum, rather than the velocity as in Eq. 2.13 or Eq. 2.13′. From Eq. 2.8 and Eq. 2.9 we get

$$E = p^0 c + \phi; \qquad m_0{}^2 c^2 = \frac{1}{c^2}(E - \phi)^2 - \mathbf{p}^2, \tag{2.15}$$

for the type of conservative potentials we are considering.

In particular, a *free particle* $[\phi = 0$, or $K = 0$ in Eq. $2.10]$ is characterized by a *constant* time-like vector, momentum vector p, with components

$$p^\mu : \left(\frac{E}{c}, \mathbf{p}\right); \qquad p^2 = m_0^2 c^2 = \frac{E^2}{c^2} = \mathbf{p}^2,$$

or,

$$E^2 = c^2 \mathbf{p}^2 + m_0^2 c^4. \tag{2.16}$$

Using the definition, Eq. 2.14, of the nonrelativistic energy and expanding E with respect to $p^2/m_0^2c^2$ in Eq. 2.15,

$$E_{NR} = \frac{1}{2}\frac{\mathbf{p}^2}{m_0} - \frac{1}{8}\frac{\mathbf{p}^4}{m_0^3 c^2} + \cdots + \phi, \tag{2.17}$$

which is the classical energy relationship with the first-order correction term in kinetic energy.

Particles with Zero Rest Mass

The concept of a zero-mass particle does not exist in Newtonian mechanics. As a matter of fact, our definition of the four-momentum, Eq. 2.7, breaks down if $m_0 = 0$. However, if we take E and $\mathbf{p}$ as primary concepts rather than the velocity or the trajectory, Eq. 2.16 still makes sense when $m_0 = 0$. In this case p cannot be connected to the velocities; any localization of the particle is not defined, or not possible. We have then a nonzero, but light-like momentum vector $p:(E/c, \mathbf{p})$, with

$$E = pc, \tag{2.18}$$

where $p_{,} = (\mathbf{p}^2)^{1/2}$. Furthermore, we see from Eq. 2.17 that there is no nonrelativistic limit to this concept. A particle with $m_0 = 0$ must be necessarily relativistic. We also note that at higher and higher momenta the particle behaves more and more like a particle with rest mass zero.

At this point we shall discuss an important realization of a particle with $m_0 = 0$, the *photon*.[1] We have seen in Chapter I, Section 4C that a plane electromagnetic wave is characterized by a light-like four-vector k (plus the polarization property), $k = (\omega/c, \mathbf{k})$, where ω is the angular frequency of the wave, and $\mathbf{k}$ is its propagation vector such that

---

[1] The concept of photon goes back to Einstein (*Ann. Phys.* *17* (1905), 132; *Physikalische Zeits.* ***10*** (1909), 185 and 817).

$$\omega = kc,$$

where $k = (\mathbf{k}^2)^{1/2}$. In various experiments, such as those concerned with blackbody radiation, the photo effect, the maximum frequency of x-ray spectra, etc., one can associate to a light plane wave an energy proportional to ω, and a momentum proportional to $\mathbf{k}$. The proportionality constants in both cases must have the same dimensions and their numerical value turns out to be equal,

$$E = \hbar\omega, \qquad \mathbf{p} = \hbar\mathbf{k}.$$

This is essential, for then we see that E and $\mathbf{p}$ defined for the plane wave exactly satisfy the energy and momentum relationship for a relativistic particle of $m_0 = 0$. In dealing with questions regarding energy and momentum interchange of light waves, say with other particles, the plane wave can be considered as a particle with $m_0 = 0$. The appearance of $\hbar$ (whose value is taken from experiment) indicates the so-called quantum-mechanical effect, and the above equations, which connect particle concepts like E and $\mathbf{p}$ with the wave concepts ω, $\mathbf{k}$, constitute a fundamental principle of quantum mechanics.[1] Another mass-zero particle is the neutrino, which is associated with the plane wave of a spinor field.

Any generalization of a special physical theory must be verified experimentally. With respect to our Minkowski equation, $\dot{p} = k$, we can state that both of its conclusions, namely the three-dimensional equation of motion or the variable mass, and the energy expression and its conservation, have been verified experimentally, the latter in nuclear-reaction experiments, where the rest masses of the particles undergo a change, and only the total energy defined with the rest energy term is conserved. Therefore, we take the Minkowski equation as the proper generalization of Newton's equations.

[1] It is interesting to indicate that, conversely, to a particle with mass m_0 one can associate a plane wave, a so-called matter wave, whose ω and $\mathbf{k}$ are again obtained by $E = \hbar\omega$, $\mathbf{p} = \hbar k$. But now k is no longer a light-like vector. From Eq. 2.16 we obtain

$$\omega^2 = c^2k^2 + \frac{m_0{}^2c^4}{\hbar^2}.$$

This relation for the de Broglie waves has been verified experimentally. L. de Broglie, *Annales de physique*, 10 serie, *2* (1925), 22.

Explicit Forms of the Minkowski Force K^μ

So far we have left the force in our fundamental equation, Eq. 2.10, arbitrary, except for the requirement that

$$K_\mu u^\mu = 0, \quad \text{or} \quad K^0 = \frac{1}{c}\,\mathbf{K}\cdot\mathbf{v}. \tag{2.11$'$}$$

This requirement does not determine K uniquely. We now discuss the possible forms of K_μ, some of which are realized in nature.

Equation 2.11′ implies immediately that some or all components of K are velocity-dependent. For example, a constant Newtonian force $\mathbf{F}$ corresponds to a Minkowski force $\mathbf{K}(v) = \mathbf{F}/c(1-\beta^2)^{1/2}$, which is increasing with velocity with $K^0 = \mathbf{F}\cdot\mathbf{v}/c^2(1-\beta^2)^{1/2}$. Even if $\mathbf{K}$ were made constant, K^0 would be velocity-dependent and *vice versa*. Because $u^\mu u_\mu = 1$, the simplest form of K_μ is a linear combination of the components of u. For this "linear" force, we write

$$K_\mu = \lambda F_{\mu\nu}(x)u^\nu. \tag{2.19}$$

The form, Eq. 2.19, will satisfy $K_\mu u^\mu = 0$ for all values of u^μ if, and only if,

$$F_{\mu\nu}(x) = -F_{\nu\mu}(x); \tag{2.20}$$

in particular, $F_{\mu\nu}$ may be constant numbers.

The electromagnetic forces on charged particles belong to this class. To see this, let us choose $\lambda = e/c$, where e is the electric charge, and connect the antisymmetric tensor $F_{\mu\nu}(x)$ to the components of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ in the following form:

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ & 0 & -B_3 & B_2 \\ & & 0 & -B_1 \\ & & & 0 \end{pmatrix}, \quad F^\mu{}_\nu = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ & 0 & B_3 & -B_2 \\ & & 0 & B_1 \\ & & & 0 \end{pmatrix};$$

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ & 0 & -B_3 & B_2 \\ & & 0 & -B_1 \\ & & & 0 \end{pmatrix}. \tag{2.21}$$

Also, we wrote the form $F^\mu{}_\nu$ and $F^{\mu\nu}$ for the purpose of convenient reference;

they follow immediately from $F_{\mu\nu}$ with our metric $g_{\mu\nu}$. The Minkowski equation with Eqs. 2.19; i.e.,

$$\frac{d}{ds}\left(m_0 c\,\frac{dx^\mu}{ds}\right) = \frac{e}{c}\,F^\mu{}_\nu(x)u^\nu, \tag{2.22}$$

are then equivalent to the equations

$$\frac{d}{dt}(m\mathbf{v}) = e\mathbf{E} + \frac{e}{c}(\mathbf{v}\times\mathbf{B}), \tag{2.22$'$}$$

and

$$\frac{d}{dt}(mc^2) = e\mathbf{E}\cdot\mathbf{v} = \frac{dT}{dt} = \frac{d}{dt}(mc^2 - m_0c^2),$$

or,

$$mc^2 = e\int \mathbf{E}\cdot d\mathbf{r}. \tag{2.22$''$}$$

Equation 2.22′ gives the Lorentz force on a particle from a given external electromagnetic field, and Eq. 2.22″ the total energy of the particle to which only the electric field contributes. Equation 2.19 indicates in addition that $F_{\mu\nu}(x)$ must transform as a tensor under Lorentz transformations if K^μ and u^μ are four-vectors. That is to say

$$F'^{\mu\nu}(x) = L^\mu{}_\sigma L^\nu{}_\rho F^{\sigma\rho}(x). \tag{2.23}$$

This gives us the transformation properties of the electric and magnetic fields; for example,

$$E_1'(x') = -L^0_\sigma L^1_\rho F^{\sigma\rho}(x),$$

and similarly for the other components. Thus, a given external electromagnetic field can be put in covariant form. The covariant form of the field equations will be discussed in the next chapter. Here we have considered $F_{\mu\nu}(x)$ from the point of view of an external force.

Next we look at the form of the Minkowski force quadratic in velocities:[1]

$$K^\mu = G^{\mu\nu\sigma}(x)u_\nu u_\sigma. \tag{2.24}$$

To satisfy the condition $K^\mu u_\mu = 0$, $G^{\mu\nu\sigma}$ must be antisymmetric in any two indices. Consequently, it will have 24 independent components. As a

---

[1] The equation of the motion of a particle in a given gravitational field is of this form. Let $g_{\mu\nu}$ be the gravitational potential. The gravitational field is given by

$$\Gamma_{\nu\sigma}{}^\mu = \tfrac{1}{2}g^{\mu\varrho}(g_{\sigma\varrho,\nu} + g_{\nu\sigma,\varrho} - g_{\nu\sigma,\varrho});$$

then the equation of motion is

$$\frac{du^\mu}{ds} = \Gamma_{\nu\sigma}{}^\mu u^\nu u^\sigma.$$

special case, we can take $G^{\mu\nu\sigma}$ to be antisymmetric in all three indices, then it will have only four independent components and can be written in the form

$$G^{\mu\nu\sigma} = \varepsilon^{\mu\nu\sigma\rho} G_{\rho}, \tag{2.24$'$}$$

where $\varepsilon^{\mu\nu\sigma\rho}$ is the completely antisymmetric symbol with elements $+1$ or -1, depending on whether $(\mu\nu\sigma\rho)$ is an even or odd permutation of $(0, 1, 2, 3)$; hence $\varepsilon^2_{\mu\nu\sigma\rho} = 4!$. G_ρ is called the dual of $G^{\mu\nu\sigma}$.

This process of building up K^μ from linear combinations of u^μ obviously can be continued to higher-order tensors.

Another class of possible force terms can be obtained by noticing that $(g_{\mu\nu} - u_\mu u_\nu)u^\nu = 0$. Hence we can put[1]

$$K^\mu = \phi^\nu(x)(g^\mu_\nu - u^\mu u_\nu). \tag{2.25}$$

Here again K^μ is determined by four functions $\phi^\nu(x)$, as in Eq. 2.24′, in contrast to the six independent components of Eq. 2.19. *In particular*, $\phi_\nu(x)$ can be thought of as the gradient of a potential

$$K^\mu = +\frac{1}{c}\frac{\partial\phi}{\partial x^\nu}(g^{\mu\nu} - u^\mu u^\nu) - \frac{1}{c}\phi\dot{u}^\mu, \tag{2.26}$$

where we have also added a term proportional to the acceleration, which we can do because $\dot{u}^\mu u_\mu = 0$. We rewrite Eq. 2.26 as

$$cK^\mu = +\frac{\partial\phi}{\partial x_\mu} - u^\mu\frac{d\phi}{ds} - \phi\dot{u}^\mu = +\phi^{,\mu} - \frac{d}{ds}(\phi u^\mu). \tag{2.26$'$}$$

For this case the Minkowski equation takes the form

$$\frac{d}{ds}\left\{\left[m_0 c + \frac{\phi(x)}{c}\right]u^\mu\right\} = +\phi^{,\mu}\frac{1}{c}. \tag{2.27}$$

In particular, if $\partial\phi/\partial x_0 = 0$ in some frame, from Eq. 2.27 we get the conservation equation

$$[m_0 c^2 + \phi(x)]u^0 = \text{const} = E, \tag{2$\cdot$28}$$

which is again our Eq. 2.13; i.e.,

$$E = mc^2 + \frac{\phi(x)}{(1-\beta^2)^{1/2}}. \tag{2.28$'$}$$

---

[1] Writing $K_\mu = A_\mu + u_\mu B^\nu u_\nu$, we find $A_\mu = -B_\mu = \phi_\mu$. This procedure can be generalized to higher tensors also: Putting $K^\mu = F^{\mu\nu}u_\nu + u^\mu u^\sigma u^\nu G_{\sigma\nu}$, we find $F^{\mu\nu} = -G^{\mu\nu}$, or

$$K^\mu = F_\sigma{}^\nu(g^{\mu\sigma}u^\nu - u^\mu u^\sigma u^\nu), \tag{2.25$'$}$$

and so on.

Thus, the potential defined in Eq. 2.26 can be considered as the relativistic generalization of the Newtonian potential. We get a potential that enters into the Newtonian energy equation only in the limit $v/c \to 0$.

Finally, we remark that if we had a constant of motion $V(x)$ along the trajectory, i.e., $dV/ds = 0$, then $K_\mu = V_{,\mu}$ would be a possible form of the Minkowski force, for $K_\mu u^\mu = V_{,\mu} u^\mu = dV/ds = 0$.

We have discussed some of the possible forms of the relativistic force and shown how to relate the origin of the force to the coupling of the velocity of the particle with external scalar, vector and tensor fields. In particular, we have introduced forces due to the external electromagnetic and scalar fields. A more general discussion of the coupling between particles and fields, which is the main subject of this book, will be given in Part II.

Angular Momentum

As in classical mechanics, the equations of motion can be expressed in terms of angular momentum. The correct generalization of the three-dimensional axial vector $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ is the second-order antisymmetric tensor, as we have seen in Chapter I in the generalization of the curl. Hence we define the covariant angular momentum by

$$J^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu, \qquad J^{\mu\nu} = -J^{\nu\mu}, \tag{2.29}$$

or, in component form,

$$J^{12} = J_z, \qquad J^{13} = -J_y, \qquad J^{23} = J_x,$$

and

$$J^{0k} = x^0 p^k - x^k p^0 = ctmv^k - x^k mc = mct\left(v^k - \frac{x^k}{t}\right).$$

We see, therefore, that the three spatial components of $J^{\mu\nu}$ represent the total angular momentum of the system $\mathbf{J}$. The remaining three time components J^{0k} do not have a simple physical meaning for a single particle. For a linear motion these components clearly vanish. In general, we can write

$$x^k = tv^k + \frac{cJ^{k0}}{mc^2}, \tag{2.30}$$

and associate J^{k0} with the deviation from the linearity of the motion. The magnitude of the angular momentum is

$$J^{\mu\nu} J_{\mu\nu} = 2[x^2 p^2 - (xp)^2] = -2J^{0k^2} + 2\mathbf{J}^2.$$

We also define a relativistic torque

$$N^{\mu\nu} = x^\mu K^\nu - x^\nu K^\mu; \tag{2.31}$$

then the equations of motion can be written in the form

$$\dot{J}^{\mu\nu} = N^{\mu\nu}, \tag{2.32}$$

using Eq. 2.7 for the linear momentum. We have here six equations instead of the four Minkowski equations, three being connected with the new quantities J^{k0}. If $N^{\mu\nu} = 0$, we get the conservation of the angular momentum $J^{\mu\nu} = \text{const}$. However, the equations $N_{0k} = 0$ do not correspond to any simple physical situation, in general. We shall come back to the discussion of the angular momentum in more specific physical problems.

Systems of Colliding Particles

So far we have considered a single particle in an external force field. The invariant formulation of the motion of two or more *interacting* particles is complicated by the fact that each particle will have a different proper time. Hence, we must use either a many-proper-time formalism, or a single invariant parameter τ for all the particles. In the latter case, simple relations like $u^2 = 1$, $p^2 = m_0^2c^2$, $\cdots$ will no longer be valid. No exact general theory seems to be available. This is one of the reasons for the significance and advance of field theory. The action-at-a-distance formulation of electrodynamics is an exception and will be discussed in Part II.

However, there is a simple situation of considerable practical interest which can be treated partly (apart from the trivial case of noninteracting particles). This is the case of collision and emission of a number of particles at a single point. If the particles originate from a point, then there is a common frame of reference in which all particles are instantaneously at rest. In that frame and at that time, $\mathbf{v} = 0$, but $d\mathbf{v}/dt \neq 0$; hence the time component of the force K^μ, K^0, is zero, and $\mathbf{K} = \mathbf{F}$, all at a given instant. Therefore, for $\mathbf{K}$ we can use the principle of action and reaction of Newtonian mechanics. Thus, summed over the momenta of all the particles, we get

$$\sum_\alpha p_\alpha = \text{const}. \tag{2.33}$$

This equation is then true in all frames because it is true in one.

In a reaction involving incoming and outgoing particles, we have to consider the sign of the momentum vectors

$$\sum_{\alpha} p_{\alpha(\text{initial})} = \sum_{\beta} p_{\beta(\text{final})} \cdot \tag{2.33$'$}$$

This is the law of the conservation of energy and momentum.[1] In components,

$$\sum_{\alpha} (\mathbf{p}_{\alpha}^2 + m_{\alpha}^2 c^2)^{1/2} = \sum_{\beta} (\mathbf{p}_{\beta}^2 + m_{\beta}^2 c^2)^{1/2}; \qquad \sum_{\alpha} \mathbf{p}_{\alpha} = \sum_{\beta} \mathbf{p}_{\beta} \cdot$$

Also it will be useful to introduce the center-of-mass frame of two (interacting) particles. This is the special frame in which the total momentum vanishes[2]

$$\mathbf{p}_1 + \mathbf{p}_2 = 0; \qquad -\mathbf{p}_2 = \mathbf{p}_1 = \mathbf{p}. \tag{2.34}$$

The center-of-mass frame momenta have the components

$$p_1 : (E_1, +\mathbf{p}), \qquad p_2 = (E_2, -\mathbf{p}).$$

From any arbitrary frame in which the momenta are p_1', p_2' (say, the laboratory frame), one can pass to the center-of-mass frame by a Lorentz transformation L defined by

$$L(p_1' + p_2') = p_1 + p_2 \,.$$

In calculating energy and momentum relations, it is convenient to make use of the fact that the scalar products $p_i p_j$ of four-vectors are Lorentz scalars and can be evaluated in any frame. For example, in the reaction $a \to b + c$, we get $m_b^2 = p_b^2 = (p_a - p_c)^2 = m_a^2 + m_c^2 - 2p_a p_c$, where the last product is equal to $m_a E_c$ in the frame in which a is at rest. Or we can form $p_a^2 = p_b^2 + p_c^2 + 2p_c p_b$, where the last product in the center-of-mass frame is $2(E_b E_c + \mathbf{p}^2)$, and in the laboratory frame is

$$2[E_b E_c - (\mathbf{p}_c)(\mathbf{p}_b)].$$

Similarly, for the reaction $a + b \to c + d$ one can form invariant relations of the form

$$(p_a + p_b)^2 = (p_c + p_d)^2,$$

or products like[3]

$$p_a(p_a + p_b) = p_a(p_c + p_d).$$

---

[1] In field theory the conservation of momentum will be related to translational invariance.

[2] One can also write a covariant definition,

$$p_1 + p_2 = 0,$$

instead of Eq. 2.34, which implies Eq. 2.34 but not *vice versa*.

[3] See Exercise 3.

2. *Lagrangian Form of the Equations of Motion*

In this and the following sections we shall discuss the generalization of the methods of analytical mechanics to the relativistic case, mainly the Lagrangian and canonical formalisms for the equations of motion. These methods do not use the concept of force and differential dynamical principles (dynamics in the small); instead they use the integral principle for the whole of the motion (dynamics in the large). In spite of their formal beauty, the integral principles seem to exclude some interesting physical situations, such as the frictional forces.

First, we wish to write a covariant Hamilton's principle for a single particle. Using the proper time introduced in the previous section, the action principle would be the following. There exists a functional of $L(x^\mu(s), \dot{x}^\mu(s), s)$ such that the true trajectory $x^\mu(s)$ makes the action integral

$$\int_{s_1}^{s_2} L(x^\mu(s), \dot{x}^\mu(s), s)\, ds$$

an extremum. More precisely, for the true motion between s_1 and s_2, the variation

$$\delta \int_{s_1}^{s_2} L(x^\mu(s), \dot{x}^\mu(s), s)\, ds = 0, \tag{2.35}$$

where the variation δ is so defined that the end points $x^\mu(s_1)$ and $x^\mu(s_2)$ are fixed, and $\delta_s = 0$ at these points. However, we cannot write formally the Euler-Lagrange equations corresponding to Eq. 2.35 because we have the subsidiary condition

$$u^2(s) = \dot{x}^{\mu^2} = 1;$$

consequently the variations are restricted. In such situations one can use either the method of Lagrange multipliers, or use an arbitrary invariant parameter τ. We discuss the latter method first.

The action principle is now

$$\delta \int_{\tau_1}^{\tau_2} L(x^\mu(\tau), v^\mu(\tau), \tau)\, d\tau = 0, \tag{2.36}$$

where we have put

$$v^\mu = \frac{dx^\mu(\tau)}{d\tau} = \frac{dx^\mu}{ds}\frac{ds}{d\tau} = u^\mu \frac{ds}{d\tau}$$

to distinguish it from u^μ, which is always understood to be the velocity with respect to the proper time s. We have also the relations

$$v^2 = \frac{dx^{\mu 2}}{d\tau^2} = \frac{ds^2}{d\tau^2} = \left(\frac{ds}{d\tau}\right)^2;$$

hence,

$$\frac{ds}{d\tau} = (v^2)^{1/2}, \quad \text{and} \quad v^\mu = u^\mu (v^2)^{1/2}.$$

It is important to realize that the parameter τ in Eq. 2.36 does not refer to a specific particular choice $\tau = \tau(s)$, but to any of the totality of such parameters. Equation 2.36 is supposed to be valid for any parameter τ, for if we fix the function $\tau = \tau(s)$, we will still get a subsidiary condition to the action principle which appears to be an additional restriction.

First, let us calculate the variation of the integrand of Eq. 2.36. We get

$$\delta L = \frac{\partial L}{\partial x^\mu}\,\delta x^\mu + \frac{\partial L}{\partial v^\mu}\,\delta v^\mu,$$

with $\delta\tau = 0$, or

$$\delta L = \left[\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial v^\mu}\right)\right]\delta x^\mu + \frac{d}{d\tau}\left(\frac{\partial L}{\partial v^\mu}\,\delta x^\mu\right), \tag{2.37}$$

where we have made use of the relation

$$\frac{d}{d\tau}(\delta x) = \delta x^\mu \left(\frac{dx}{d\tau}\right)$$

For the particular variation indicated in Eq. 2.25 or 2.36; i.e., with end points fixed and $\delta\tau = 0$, we have

$$\delta \int L\, d\tau = \int \delta L\, d\tau.$$

In the integration of δL, given by Eq. 2.37, the second term does not contribute because δx^μ is fixed and zero at the boundaries:

$$\delta \int L\, d\tau = \int \left[\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial v^\mu}\right)\right]\delta x^\mu\, d\tau = 0.$$

And, because δx^μ is arbitrary inside the boundaries, we get the equation of motion.

$$\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial v^\mu}\right) = 0, \tag{2.38}$$

which must be understood to hold for an arbitrary unspecified parameter τ.

When the Euler-Lagrange equations, Eq. 2.38, hold, we can write a posteriori for the variation of L from Eq. 2.37:

$$\delta L = \frac{d}{d\tau}\left(\frac{\partial L}{\partial v^\mu}\,\delta x^\mu\right), \tag{2.39}$$

with $\delta\tau = 0$.

Examples: (a) A free particle may be defined by the Lagrangian

$$L_0 = m_0 c (v^2)^{1/2} = m_0 c \frac{ds}{d\tau}. \tag{2.40}$$

Then Eq. 2.39 gives

$$\frac{d}{d\tau}\left[m_0 c \frac{v^\mu}{(v^2)^{1/2}}\right] = 0,$$

or

$$\frac{d}{ds}(m_0 c u^\mu) = 0,$$

which is the Minkowski equation for a free particle.[1]

(b) Consider the Lagrangian

$$L = L_0 + L_1 = m_0 c (v^2)^{1/2} + \frac{e}{c} A_\mu(x) v^\mu, \tag{2.41}$$

where the interaction part of L couples the particle velocity to a (external) vector field $A_\mu(x)$. Eq. 2.39 gives, in this case,

$$\frac{e}{c} v^\nu \frac{\partial A_\nu}{\partial x^\mu} - \frac{d}{d\tau}\left[m_0 c \frac{v_\mu}{(v^2)^{1/2}} + \frac{e}{c} A_\mu\right] = 0,$$

$$\frac{d}{ds}(m_0 c u_\mu) = \frac{e}{c}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right) u^\nu \equiv \frac{e}{c} F_{\mu\nu} u^\nu. \tag{2.42}$$

This is the Minkowski equation, with the force term given by Eq. 2.19, that we have identified as that of a particle moving in an external electromagnetic field. We also see that the "coupling" in the Lagrangian is in terms of the so-called electromagnetic potentials $A_\mu(x)$, and that the electromagnetic field strengths are given by the curl of the potential,

$$F_{\mu\nu}(x) = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}. \tag{2.43}$$

---

[1] Note that in this example $\int L\, d\tau = \int m_0 c\, ds$. However, one cannot take $m_0 c$ as the Lagrangian, one must also vary s, whereas in our formulation we did not vary τ. Both these methods and many others give the same result in simple examples.

Writing

$$A^{\mu}:(\phi, \mathbf{A}), \qquad A_{\mu}:(\phi, -\mathbf{A}), \tag{2.44}$$

and using Eqs. 2.21, Eq. 2.43 can be written in the more familiar component forms

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla \times \mathbf{A}, \tag{2.43$'$}$$

where ϕ is the scalar potential and **A** the vector potential. Note that for **A** = constant number we get a Lorentz force

$$K^0 = -\frac{1}{c}\,\mathbf{u}\cdot\nabla\phi, \qquad \mathbf{K} = -\frac{1}{c}\,u^0\nabla\phi,$$

discussed earlier, which differs from the pure scalar field of Eq. 2.26.

(c) As a third example, we consider the Lagrangian

$$L = L_0 + L_1 = [m_0c + \phi(x)](v^2)^{1/2}, \tag{2.45}$$

which leads to the equation of motion, 2.27, discussed before.

(d) Another example of a Lagrangian homogeneous of first order in velocities and describing the interaction with an external tensor field is

$$L = \left[m_0c - \phi_{\mu\nu}\frac{v^{\mu}v^{\nu}}{(v^2)^{1/2}}\right](v^2)^{1/2}, \tag{2.45$'$}$$

which leads to the equation of motion

$$\frac{d}{ds}(m_0cu_{\mu}) = \frac{d}{ds}[(\phi_{\mu\nu} + \phi_{\nu\mu})u^{\nu} - \phi_{\sigma\rho}u^{\sigma}u^{\rho}u_{\mu}] - \phi_{\sigma\rho,\mu}u^{\sigma}u^{\rho}.$$

Generalized Momenta. Returning to our general equation, Eq. 2.38, we define the generalized momenta by

$$p_{\mu} = \frac{\partial L}{\partial v^{\mu}}, \tag{2.46}$$

which gives

$$\frac{d}{d\tau}p_{\mu} = \frac{\partial L}{\partial x^{\mu}}. \tag{2.38$'$}$$

If the Lagrangian does not contain a coordinate x^{ν} explicitly, then the corresponding component p^{ν} of the momentum is independent of τ or conserved.

For the examples discussed before, Eqs. 2.40, 2.41, 2.45, 2.45′, we obtain, respectively,

$$p^\mu = m_0 c u^\mu = \text{const},$$

$$p^\mu = m_0 c u^\mu + \frac{e}{c} A^\mu(x),$$

$$p^\mu = [m_0 c - \phi(x)] u^\mu,$$

$$p^\mu = m_0 c u_\mu - (\phi_{\mu\nu} + \phi_{\nu\mu}) u^\nu + \phi_{\sigma\rho} u^\sigma u^\rho u^\mu.$$

Example (b) shows the difference between the generalized or canonical momentum and that used in the Minkowski equation, Eq. 2.7.

Variation of Endpoints. Consider a Lagrangian not explicitly dependent on τ; i.e., $L = L(x^\mu(\tau),\ v^\mu(\tau))$, and a new kind of variation which also involves the variation of the end points,

$$\delta W = \delta \int_{\tau_1}^{\tau_2} L\, \delta\tau = \int_{\tau_1}^{\tau_2} \delta L\, d\tau + L\, \delta\tau \Big|_{\tau_1}^{\tau_2}. \tag{2.47}$$

For δL we use Eq. 2.37, assuming, however, that the equations of motion are satisfied (i.e., for the actual path); then

$$\delta W = \left[\frac{\partial L}{\partial v^\mu}\, \delta x^\mu + L\, \delta\tau\right]_{\tau_1}^{\tau_2} = F(\tau)\Big|_{\tau_1}^{\tau_2} = F(\tau_2) - F(\tau_1). \tag{2.48}$$

This equation is important in the discussion of constants of motion; namely, if we can find a variation such that W is invariant under it; i.e., $\delta W = 0$, then $F(\tau_2) = F(\tau_1)$, or $F(\tau)$ has to be a constant of motion. F is called the *generator* of the transformations on W. *The generators of invariant transformations are constants of motion.*

We now try to introduce formally a Hamiltonian, as in nonrelativistic analytical mechanics. Consider a Lagrangian $L = L(x^\mu, v^\mu)$, and form

$$\frac{dL}{d\tau} = \frac{\partial L}{\partial x^\mu} v^\mu + \frac{\partial L}{\partial v^\mu} \dot{v}^\mu.$$

Using the equations of motion, Eq. 2.38, this equation can be rewritten as

$$\frac{dL}{d\tau} = \frac{d}{d\tau}\left(v^\mu \frac{\partial L}{\partial v^\mu}\right),$$

or

$$\frac{d}{d\tau}\left(L - v^\mu \frac{\partial L}{\partial v^\mu}\right) = \frac{dC}{d\tau} = 0 \tag{2.49}$$

In the nonrelativistic case, C obtained by this procedure is the Hamiltonian and has the interpretation of energy.[1] In the relativistic case, energy is the zero component of the momentum. We shall see that C can be interpreted formally as the "Hamiltonian" [even though it is not energy and its numerical value is zero (see below)] as far as the derivation of the Hamilton equations is concerned. But other quantities can also be used as the "Hamiltonian" which *are* related to the energy.

From our requirement of parameter-invariance, Eq. 2.49 must be true for all τ. The Lagrangian L has the same form for all τ, but different numerical values. Thus, we conclude that

$$C \equiv L - v^\mu \frac{\partial L}{\partial v^\mu} = 0, \tag{2.50}$$

which means that L is a function that is *homogeneous of the first order in velocities* v^μ.[2] We have then the relation

$$L(\lambda v^\mu) = \lambda L(v^\mu). \tag{2.50$'$}$$

Now, if we write the Lagrangian as a sum of the free-particle Lagrangian L_0 plus an interaction Lagrangian L_1,

$$L = L_0 + L_1, \tag{2.51}$$

where both L_0 and L_1 are homogeneous of the first order in velocities, then the Minkowski force is given by

$$K^\mu = \frac{\partial L_1}{\partial x^\mu}\frac{d\tau}{ds} - \frac{d}{ds}\left(\frac{\partial L_1}{\partial v^\mu}\right), \tag{2.52}$$

and automatically satisfies the condition $K^\mu u_\mu = 0$. But the converse of this statement is not true. There are examples of K^μ satisfying $K^\mu u_\mu = 0$ (see Sect. 1) that are not homogeneous of the first order in velocities. Therefore, we conclude that the requirement of the parameter invariance of L invokes an additional restriction on the Minkowski equation, as we have observed before.

Method of Lagrange multipliers. We now discuss the solution of the original variational principle, Eq. 2.35, in the variable s with the subsidiary

---

[1] We recall that for $L = L(q, \dot{q})$, $dL/dt = (\partial L/\partial q)\dot{q} + (\partial L/\partial \dot{q})\ddot{q}$, or, with the use of the equations of motion,

$$\frac{dL}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\dot{q}\right), \quad \text{or} \quad \frac{d}{dt}(L - p\dot{q}) = \frac{dH}{dt} = 0.$$

[2] A function $f(x)$ homogeneous of nth order satisfies the differential equation $nf = x(\partial f/\partial x)$, and $f(\lambda x) = \lambda^n f(x)$.

condition $u^2 = 1$. Let the Lagrangian L be explicitly independent of s; i.e., $L = L[x^\mu(s), \dot{x}^\mu(s)]$. We introduce a Lagrange multiplier $\lambda/2$ and use a new Lagrangian $\bar{L}$ defined by[1]

$$\bar{L} = L + \frac{\lambda}{2}(u^2 - 1). \tag{2.53}$$

Now the variational principle $\delta \int \bar{L} ds = 0$ leads directly to the Euler-Lagrange equations,

$$\frac{\partial L}{\partial x} - \frac{d}{ds}\left(\frac{\partial L}{\partial u^\mu}\right) - \frac{d}{ds}(\lambda u_\mu) = 0. \tag{2.54}$$

The variation with respect to the multiplier λ gives, in addition,

$$u^2 - 1 = 0. \tag{2.55}$$

Now we have five equations for five unknowns $x^\mu(s)$ and $\lambda(s)$. Multiplying Eq. 2.54 with u^μ and noting that $u^2 = 1$, $u\dot{u} = 0$, we find

$$\dot{\lambda} = u^\mu \frac{\partial L}{\partial x^\mu} - u^\mu \frac{d}{ds}\left(\frac{\partial L}{\partial u^\mu}\right),$$

or

$$\begin{aligned}
\lambda = \int \dot{\lambda}\, ds &= \int \frac{\partial L}{\partial x^\mu}\frac{dx^\mu}{ds}\, ds - \int u^\mu \frac{d}{ds}\left(\frac{\partial L}{\partial u^\mu}\right) ds \\
&= \int dL - \int \frac{\partial L}{\partial u^\mu}\dot{u}^\mu\, ds - \int u^\mu \frac{d}{ds}\left(\frac{\partial L}{\partial u^\mu}\right) ds \\
&= L - \int \frac{d}{ds}\left(u^\mu \frac{\partial L}{\partial u^\mu}\right) ds \\
&= L - u^\mu \frac{\partial L}{\partial u^\mu}.
\end{aligned} \tag{2.56}$$

With this, the final form of the equation of motion is

$$\Omega_\mu \equiv \frac{\partial L}{\partial x^\mu} - \frac{d}{ds}\left(\frac{\partial L}{\partial u^\mu}\right) - \frac{d}{ds}\left[\left(L - u^\nu \frac{\partial L}{\partial u^\nu}\right)u_\mu\right] = 0. \tag{2.57}$$

The generalized momenta now must be defined by

$$\bar{p}_\mu = \frac{\partial \bar{L}}{\partial u^\mu},$$

and are given by

$$\bar{p}_\mu = \frac{\partial L}{\partial u^\mu} + \left(L - u^\nu \frac{\partial L}{\partial u^\nu}\right)u_\mu. \tag{2.58}$$

---

[1] L. Infeld, *Bull. Acad. Pol. Sci.*, 5 (1957), 491.

The equations of motion, Eq. 2.57, are indeed of the form

$$\dot{\bar{p}}_\mu = \frac{\partial L}{\partial x^\mu}. \tag{2.57'}$$

One can verify easily that for the three typical examples discussed in Eqs. (2.40)ff, the Lagrangians to be used in connection with the equations of motion, Eq. 2.57, are

$$L = m_0 c,$$

for a free particle,

$$L = m_0 c + \frac{e}{c} A_\mu u^\mu,$$

particle in an external electromagnetic field,

$$L = m_0 c - \phi(x), \tag{2.59}$$

particle in an external scalar field.
The generalized momenta are exactly the same as before.

Lagrangian containing the subsidiary condition $u^2 = 1$. Finally, we discuss the possibility that one can choose a Lagrangian $L = L(x^\mu, u^\mu)$ in such a way that the condition $u^2 = 1$ is satisfied automatically by virtue of the equations of motion. Neglecting for the moment the condition $u^2 = 1$, we obtain from Eq. 2.36

$$\frac{\partial L}{\partial x^\mu} = \dot{p}_\mu, \qquad p_\mu = \frac{\partial L}{\partial u^\mu} \tag{2.60}$$

provided we can choose subsequently the correct Lagrangian to satisfy the condition $u^2 = 1$. That this can be done is shown by the observation that, in the three typical examples considered, had we used the parameter s rather than τ, we would have obtained the correct equations of motion; hence, $u^2 = 1$ must be automatically satisfied. More precisely the Lagrangians

$$L = m_0 c u^2,$$

$$L = m_0 c u^2 + \frac{e}{c} A_\mu u^\mu, \tag{2.61}$$

$$L = (m_0 c - \phi) u^2,$$

give the correct equations of motion, from which we obtain

$$K^\mu u_\mu = \dot{u}_\mu u^\mu = 0,$$

or $u_\mu u^\mu = \text{const}$, which can be chosen equal to 1. The point here is to keep the u^2 in L as far as the functional dependence of L on u^μ is concerned.[1]

3. *Canonical Form of the Equations of Motion*

We recall that in classical mechanics the n Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = \frac{\partial L}{\partial q_k}, \qquad k = 1, 2, \cdots, n,$$

for the n coordinate functions $\dot{q}_k = \dot{q}_k(t)$, can be transformed into $2n$ Hamilton equations as follows. One defines the generalized momenta

$$p_k = \frac{\partial L}{\partial \dot{q}_k},$$

and solves this equation for $\dot{q}_k = \dot{q}_k(q_k, p_k, t)$. If this can be done, one can express the Hamiltonian function

$$H = p_k \dot{q}_k - L$$

as a function of q_k and p_k: $H = H(q_k, p_k, t)$. We then obtain the canonical equations

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}; \qquad \dot{q}_k = \frac{\partial H}{\partial p_k}.$$

In this section we discuss the corresponding equations for the relativistic case.

We have already introduced the generalized momenta and noticed that a straightforward construction of the "Hamiltonian" for a single particle from the Lagrangian leads to an expression that is identically zero.

$$\mathscr{H} = u^\mu p_\mu - L = 0. \tag{2.62}$$

However, we can still use $\mathscr{H}$ to derive the canonical equations by differentiating with respect to its functional dependence. For this purpose $\mathscr{H}$ must be expressed as a function of x^μ and p^μ, which we now take as independent variables.

$$\mathscr{H} = \mathscr{H}(x^\mu, p^\mu). \tag{2.62$'$}$$

---

[1] A further form of the action principle as an integral in four-dimensional space will be discussed in detail in Part II. This form will be suitable when we consider particles and fields as two physical systems in their own right and in their mutual interaction.

Using the equations of motion $\dot{p}_\mu = \partial L/\partial x^\mu$, we obtain the canonical equations formally from Eq. 2.62:

$$\begin{aligned} \frac{\partial \mathscr{H}}{\partial x^\mu} &= -\dot{p}_\mu, \\ \frac{\partial H}{\partial p^\mu} &= \dot{x}_\mu. \end{aligned} \tag{2.63}$$

For a general Lagrangian combining our three typical examples; i.e., $L = [m_0 c - \phi(x)]u^2 + (e/c)A_\mu u^\mu$, we get

$$\mathscr{H} = u^\mu p_\mu - [m_0 c - \phi(x)]u^2 - \frac{e}{c} A_\mu u^\mu,$$

which must be considered as expressed as a function of x^μ and p_μ. Then

$$\frac{\partial \mathscr{H}}{\partial x^\mu} = \phi_{,\mu} u^2 - \frac{e}{c}\frac{\partial A_\nu}{\partial x^\mu} u^\nu = -\dot{p}_\mu; \qquad \frac{\partial \mathscr{H}}{\partial p^\mu} = u_\mu = \dot{x}_\mu.$$

Although this method is formally correct, one can argue against it on two grounds; first, $\mathscr{H}$ is identically vanishing, and secondly, it does not have the meaning of energy. To avoid the first difficulty, one can use instead of $\mathscr{H}$, the symmetric tensor

$$H^{\mu\nu} = p^\mu u^\nu - g^{\mu\nu} L, \tag{2.64}$$

with the canonical equations

$$\frac{\partial H^{\mu\nu}}{\partial x^\mu} = -\dot{p}^\nu; \qquad \frac{\partial H^{\mu\nu}}{\partial p^\nu} = u^\nu. \tag{2.65}$$

To obtain a Hamiltonian that has the meaning of energy, we must proceed differently.

We have seen that, in general, the canonical momentum p^μ consists of a kinetic term $m_0 c u^\mu$, the momentum of a free particle, plus a term independent of the rest mass m_0; i.e.,

$$p^\mu = m_0 c u^\mu + f^\mu(\phi \cdots, u), \tag{2.66}$$

where f^μ is a function of the external fields and velocities. Because $u^2 = 1$, we get

$$(p - f)^2 = m_0^2 c^2, \tag{2.67}$$

which generalizes the free-particle equation $p^2 = m_0c^2$, Eq. 2.9.[1] Therefore, the quantity $(p^\mu - f^\mu)$ is the relevant and important quantity as far as the motion is concerned. The equation of motion can be written in the form

$$\frac{d}{ds}(p^\mu - f^\mu) = \frac{d}{ds}(m_0cu^\mu) = \frac{\partial L}{\partial x_\mu} = K^\mu. \tag{2.68}$$

In this form it is the same as the Minkowski equation (Eq. 2.10), written now in terms of the canonical momentum p^μ, whereas in Eq. 2.10 we wrote the m_0cu^μ as the kinetic momentum p^μ. Now, of course, the energy and momentum of the particle will be different from that of a free particle. The energy of the interacting particle is obtained from Eq. 2.67 to be (assuming f is independent of velocities)

$$E = cp^0 = cf^0 + c[(\mathbf{p} - \mathbf{f})^2 + m_0^2c^2]^{1/2} = H. \tag{2.69}$$

If f^μ depends on u^μ, as in the example of the tensor field, Eq. 2.45′, we have to re-express u^μ in terms of the p^μ in Eq. 2.67. Now E in Eq. 2.69 can be used as the Hamiltonian to write the canonical equations as in classical mechanics,

$$\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{x}}, \qquad \frac{\partial H}{\partial \mathbf{x}} = -\dot{\mathbf{p}}. \tag{2.70}$$

Equations 2.70 are not covariant, but they are perfectly relativistically invariant. We have already pointed out that covariance is not necessary (it is sufficient) for the relativistic invariance. Although H does not transform as a scalar, vector, tensor or spinor, nevertheless it has a definite transformation property, and Eqs. 2.70 can be written down in every frame. H transforms like the zero component of a four-vector.

We leave the discussion of our examples from the point of view of Eqs. 2.70 to the exercises at the end of this chapter. We note here that Eq. 2.69 for small v/c becomes

$$E - m_0c^2 = E_{NR} = cf^0 + \frac{1}{2m_0}(\mathbf{p} - \mathbf{f})^2 - \frac{1}{8m_0^3c^2}(\mathbf{p} - \mathbf{f})^4 + \cdots. \tag{2.71}$$

---

[1] In the case of a scalar external field, we find, from $p^\mu = m_0cu^\mu - \phi u^\mu$,

$$u^\mu = p^\mu \frac{1}{m_0c - \phi};$$

hence

$$p^2\left(1 + \frac{\phi}{m_0c - \phi}\right)^2 = m_0{}^2c^2,$$

or

$$p^2 = p^{02} - \mathbf{p}^2 = (m_0c - \phi)^2; \qquad E = c[\mathbf{p}^2 + (m_0c - \phi)^2]^{1/2}.$$

We also find that

$$(p^2)^{1/2} + \phi = m_0c.$$

For example, in the case of an external electromagnetic field, $f^\mu = (e/c)A^\mu$, and we get the more familiar equations

$$E = e\phi + c\left[\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + m_0^2c^2\right]^{1/2}; \tag{2.72}$$

$$E_{NR} = e\phi + \frac{1}{2m_0}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 - \frac{1}{8m_0^3c^2}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^4 + \cdots. \tag{2.73}$$

It is still possible to write covariant canonical equations with a "Hamiltonian" that is a Lorentz scalar and that is related to the rest energy of the particle. Consider the following two scalars:

$$M = \frac{1}{2m_0c}(p - f)^2 = \tfrac{1}{2}m_0c; \tag{2.74}$$

$$M' = [(p - f)^2]^{1/2} = m_0c, \tag{2.75}$$

where f is independent of u^μ. From Eq. 2.74 and Eq. 2.66, we find

$$\frac{\partial M}{\partial p^\mu} = \frac{1}{m_0c}(p_\mu - f_\mu) = u_\mu = \dot{x}_\mu;$$

$$\frac{\partial M}{\partial x^\mu} = -\frac{1}{m_0c}(p^\nu - f^\nu)f_{\nu,\mu} = -u^\nu f_{\nu,\mu} = -\dot{p}_\mu;$$

i.e., correct canonical equations. Similar equations hold for M'. From $M'^2 = (p - f)^2$, we get

$$M'\frac{\partial M'}{\partial x^\mu} = (p^\nu - f^\nu)f_{\nu,\mu}; \qquad M'\frac{\partial M'}{\partial p^\mu} = (p_\mu - f_\mu); \tag{2.76}$$

using the value of $M' = m_0c$, we get back the previous canonical equations.

With respect to M', we note that in going over to quantum theory the operation of taking the square root of the operators is not straightforward and leads to the introduction of an additional finite-dimensional internal space, as in the case of the Dirac theory.

For a scalar field, the functions M and M' are given by

$$M = \frac{1}{2m_0c}[(p^2)^{1/2} + \phi]^2 = \tfrac{1}{2}m_0c; \qquad M' = (p^2)^{1/2} + \phi = m_0c, \tag{2.77}$$

respectively.

To summarize our considerations on Lagrangian and canonical forms, we can state that there are a multitude of forms in both cases that are essentially equivalent, and that the problem of writing relativistic Lagrangian and canonical equations poses no great difficulty.

Also, the development of canonical transformations, Poisson brackets, and the Hamilton-Jacoby theory can be carried over formally to the relativistic case with any of the "Hamiltonian" functions we have discussed above. For completeness, we give a brief discussion of these topics.

If one uses the noncovariant relativistic Hamiltonian, Eq. 2.69, the discussion is exactly the same as in the case of analytical mechanics and need not be repeated here.[1] Therefore, we shall comment on the covariant forms only.

Consider our scalar Hamiltonian M or M', or H, which we denote by $\mathscr{H}$ in the following, meaning any one of them. Introducing an eight-dimensional phase space (p^μ, x^μ) the *canonical transformations* may be defined by the equation

$$p_\mu v^\mu - \mathscr{H} = p'_\mu v'^\mu - \mathscr{H}' + \frac{dF}{ds}, \tag{2.78}$$

where F is the generator of the canonical transformation. Choosing, for example, $F = F(x^\mu, x'^\mu, s)$ (there are three other forms possible, F must be a function of one set of all coordinates and one set of new coordinates), we obtain

$$\begin{aligned} H' &= H + \frac{\partial F}{\partial S}, \\ p_\mu &= \frac{\partial F}{\partial x^\mu} \\ p'_\mu &= -\frac{\partial F}{\partial x'^\mu} \end{aligned} \tag{2.79}$$

The *Poisson bracket* of two functionals of x^μ and p^μ will be defined formally as

$$[P, Q] = \frac{\partial P}{\partial x^\mu}\frac{\partial Q}{\partial p_\mu} - \frac{\partial P}{\partial p^\mu}\frac{\partial Q}{\partial x_\mu}. \tag{2.80}$$

If P and Q are scalar functions, $[P, Q]$ is invariant under Lorentz transformations, and it has all the usual properties of the Poisson brackets.

In particular, using Eq. 2.63, we find

$$[x^\mu, H] = \dot{x}^\mu; \quad [p^\mu, H] = \dot{p}^\mu. \tag{2.81}$$

For a general scalar function $P(p^\mu, x^\mu)$, we obtain

$$[P, H] = \frac{dP}{ds}; \tag{2.82}$$

---

[1] See, for example, H. Goldstein, *Op. cit.*, Chapter 8.

if P also depends explicitly on s, we have the more general formula

$$[P, H] = \frac{dP}{ds} + \frac{\partial P}{\partial s}. \tag{2.83}$$

Therefore, quantities independent of the proper time s commute with the Hamiltonian.

Thus, we see that the general theory of canonical transformations and Poisson brackets—as well as the theory of infinitesimal canonical transformations and Hamilton-Jacoby theory (see Exercise 9)—can be put in covariant form for a single particle. As we have noted before, in the case of several particles, we cannot use a common proper time. Consequently, no exact relativistic theory of several interacting particles can be given by the methods of the present chapter. (See Chapter VI.)

4. *Electric and Magnetic Moments: Classical Spin*

In this section we shall discuss some new intrinsic properties of our classical relativistic mass point which manifest themselves by additional terms in the equation of motion. So far we have only considered a mass point completely characterized by its rest mass in a given external scalar, vector or tensor field. A spinor field is included in the sense that, in general, tensors can be constructed out of spinors. The spinor formulation of mechanics itself will be discussed in the next section.

We start with the case of an external electromagnetic field. If the mass point has an intrinsic magnetic moment $\boldsymbol{\mu}$ (not the orbital magnetic moment), then it has, in an external magnetic field $\mathbf{B}$, an additional energy, $\boldsymbol{\mu}\cdot\mathbf{B}$. Similarly, a particle with an intrinsic dipole moment $\mathbf{d}$ has an additional energy $\mathbf{d}\cdot\mathbf{E}$ in an external electric field E. As a matter of fact, these moments may be defined phenomenologically by the form of the energy. The covariant generalization of these terms in the energy can be written as

$$\tfrac{1}{2}\sigma^{\mu\nu}F_{\mu\nu}, \tag{2.84}$$

where $\sigma^{\mu\nu}$ is an antisymmetric tensor:

$$\sigma^{\mu\nu} = \begin{pmatrix} 0 & d_1 & d_2 & d_3 \\ -d_1 & 0 & \mu_3 & -\mu_2 \\ -d_2 & -\mu_3 & 0 & \mu_1 \\ -d_3 & \mu_2 & -\mu_1 & 0 \end{pmatrix}. \tag{2.85}$$

In the same way as electric and magnetic fields form a unity $F_{\mu\nu}$, electric and magnetic moments form a unity. In particular, these are transformed into each other under Lorentz transformations.

Because the moments are not scalars but vectors in the rest frame of the particle, their orientation will depend on the motion and, in general, will change along the trajectory. We are faced then with two problems:

(a) Determination of the motion in the presence of the moments;

(b) Determination of the precessions of the moments.

We solve the problem (a) first for a constant $\sigma_{\mu\nu}$. We may start conveniently from a Lagrangian, which must contain the additional term, Eq. 2.84:

$$L = (m_0 c + \tfrac{1}{2}\sigma^{\mu\nu}F_{\mu\nu})(v^2)^{1/2} + \frac{e}{c} A_\mu v^\mu. \tag{2.86}$$

According to Eq. 2.38 this Lagrangian leads to the equation of motion

$$\frac{d}{ds}(m_0 c u_\mu) = \frac{e}{c} F_{\mu\nu}u^\nu + \tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta,\mu} - \frac{d}{ds}(\tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta}u_\mu).$$

The last term still contains a factor *du*/*ds*, so the equation is not yet in a form to allow a physical interpretation. Rewriting the last term,

$$\frac{d}{ds}(\tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta}u_\mu) = \tfrac{1}{2}u_\nu u_\mu \sigma^{\alpha\beta}F_{\alpha\beta}{}^{;\nu} + \tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta}\frac{du_\mu}{ds},$$

and grouping terms we get

$$\frac{d}{ds}(m_0 c u_\mu) = \frac{1}{1+\Delta m}\left(\frac{e}{c} F_{\mu\nu}u^\nu + \tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta,\mu} - \tfrac{1}{2}u_\nu u_\mu \sigma^{\alpha\beta}F_{\alpha\beta}{}^{;\nu}\right), \tag{2.87}$$

where

$$\Delta m = \frac{1}{2m_0c^2}\sigma^{\alpha\beta}F_{\alpha\beta}.$$

The effect of the moment terms is therefore twofold: one, to change effectively the mass or the charge of the particle [the factor $1/(1 + \Delta m)$]; and two, to add forces which depend on the gradient of the external field. Even for a homogeneous field, we get a deviation from the Lorentz force:

$$F^{\alpha\beta}_{,\mu} = 0 : \frac{d}{ds}(m_0 c u_\mu) = \frac{1}{1+\Delta m}\frac{e}{c} F_{\mu\nu}u^\nu. \tag{2.88}$$

If the energy of the moments $\tfrac{1}{2}\sigma^{\alpha\beta}F_{\alpha\beta}$ is very small compared with the rest energy of the particle, Δm is also very small and the Lorentz equation holds.

The above equations are not very realistic, because $\sigma^{\mu\nu}$ in general is a function of s and we must consider the problem of the precession of the moments. We can use either the tensor $\sigma^{\mu\nu}$, or a vector S^μ dual to $\sigma^{\mu\nu}$ and defined by

$$\sigma_{\mu\nu} = \tfrac{1}{2}\varepsilon_{\mu\nu\sigma\rho}p^\sigma S^\rho, \tag{2.89}$$

where p^μ is the momentum vector, or

$$S^\mu = \tfrac{1}{2}\varepsilon^{\mu\nu\sigma\rho}p_\nu\sigma_{\sigma\rho}\,. \tag{2.90}$$

From Eq. 2.89 and Eq. 2.90, we obtain the relations

$$p_\mu S^\mu = 0, \qquad p_\mu\sigma^{\mu\nu} = 0 = S_\mu\sigma^{\mu\nu}, \qquad \sigma_{\mu\nu} = -\sigma_{\nu\mu}\,. \tag{2.91}$$

Thus, spin may also be described by a four-vector orthogonal to the momentum p. In the rest frame of the particle $p:(m, 0, 0, 0)$, the spin vector S is a three-dimensional space-like vector $(0, \mathbf{S})$.

During the motion of the particle, S^μ or $\sigma_{\mu\nu}$ will change. There are various methods of deriving an equation of motion for these quantities as a function of proper time, from which the precession of the spin follows immediately. Of course, the motion of the spin is such that Eqs. 2.91 are valid at all times.

One simple method is to start from the equation of motion in the rest frame of the particle.[1] In this frame, the equation of motion in external electromagnetic fields $(\mathbf{E}, \mathbf{B})$ is given by

$$\frac{d\mathbf{S}}{dt} = \left(\frac{g}{2}\right)\frac{e}{m}\mathbf{S}\times\mathbf{B} + \left(\frac{f}{2}\right)\frac{e}{m}\mathbf{S}\times\mathbf{E}, \tag{2.92}$$

where g and f measure the magnetic and the electric moments. The moments themselves are given by

$$\boldsymbol{\mu} = \left(\frac{g}{2}\right)\frac{e}{m}\mathbf{S}; \qquad \boldsymbol{\rho} = \left(\frac{f}{2}\right)\frac{e}{m}\mathbf{S}. \tag{2.93}$$

We also note that $\boldsymbol{\mu}$ and $\boldsymbol{\rho}$ as the spatial part of a vector, or as components of an antisymmetric tensor, transform as follows:

$$\begin{aligned} \boldsymbol{\mu}' &= \gamma\boldsymbol{\mu} + (1-\gamma)\frac{\mathbf{v}\cdot\boldsymbol{\mu}}{v^2}\mathbf{v} + \frac{\gamma}{c}\mathbf{v}\times\boldsymbol{\rho}; \\ \boldsymbol{\rho}' &= \gamma\boldsymbol{\rho} + (1-\gamma)\frac{\mathbf{v}\cdot\boldsymbol{\rho}}{v^2}\mathbf{v} + \frac{\gamma}{c}\mathbf{v}\times\boldsymbol{\mu}, \end{aligned} \tag{2.94}$$

where

$$\gamma = (1-\beta^2)^{-1/2}.$$

---

[1] V. Bargmann, L. Michel, and V. Telegdi, *Phys. Rev. Letters*, 2 (1959), 435. This reference deals with homogeneous fields only.

The unique generalization of Eq. 2.92 to any frame is

$$\frac{dS^\mu}{d\tau} = \left(\frac{g}{2}\right)\frac{e}{m}\,[F^{\mu\nu}u_\nu + (S_\sigma F^{\sigma\rho}u_\rho)u^\mu]$$
$$+ \left(\frac{f}{2}\right)[F^{\mu\nu *}u_\nu + (S_\sigma F^{\sigma\rho *}u_\rho)u^\mu] - \frac{du^\nu}{d\tau}\,S_\nu u^\mu, \quad \mathbf{(2.95)}$$

where $F^*_{\mu\nu}$ is the dual of $F_{\mu\nu}$ and is obtained by replacing E by $+B$ and B by $-E$, or

$$F^*_{\mu\nu} = \tfrac{1}{2}\varepsilon_{\mu\nu\sigma\rho}F^{\sigma\rho}. \quad \mathbf{(2.96)}$$

In Eq. 2.95 the factor $du^\mu/d\tau$ in the last term is given by the equation of motion, Eq. 2.87, which we have determined previously. If we substitute this equation into Eq. 2.95, we obtain the final result:

$$\frac{dS^\mu}{d\tau} = \frac{e}{m}\left[\left(\frac{g}{2}\right)F^{\mu\nu}S_\nu + \left(\frac{g}{2}\right)F^{\mu\nu}S_\nu + \left(\frac{g}{2} - \frac{1}{1+\Delta m}\right)S_\sigma F^{\sigma\rho}u_\rho u^\mu\right.$$
$$\left. + \frac{f}{2}\,F^{\mu\nu *}S_\nu + \frac{f}{2}\,S_\sigma F^{\sigma\rho *}u_\rho u^\mu\right] - \frac{S_\sigma \sigma^{\rho\nu}F^{;\sigma}_{\rho\nu}u^\mu}{2mc^2(1+\Delta m)}. \quad \mathbf{(2.97)}$$

A simpler formula is obtained in the special case of homogeneous fields and, neglecting Δm,

$$\frac{dS^\mu}{d\tau} = \frac{e}{m}\left[\frac{g}{2}\,F^{\mu\nu}S_\nu + \left(\frac{g}{2} - 1\right)S_\sigma F^{\sigma\rho}u_\rho u^\mu + \frac{f}{2}\,F^{\mu\nu *}S_\nu + \frac{f}{2}\,S_\sigma F^{\sigma\rho *}u_\sigma u^\mu\right]. \quad \mathbf{(2.98)}$$

This is the result obtained by Bargmann, Michel and Telegdi, referred to above.

Next we discuss a Hamiltonian form of the equation of motion of the spin. Consider the equation of motion

$$\frac{d\mathbf{S}}{dt} = \left(\frac{g}{2}\right)\frac{e}{m}\,\mathbf{S} \times \mathbf{B} \quad \mathbf{(2.99)}$$

of a magnetic moment in the rest frame of the particle. Because the magnitude of S, S, is constant, there are only two degrees of freedom in the problem. We choose

$$\xi = S_z, \quad \text{and} \quad \eta = \tan^{-1}(S_x/S_y) \quad \mathbf{(2.100)}$$

as the canonical variables and a Hamiltonian,

$$H = \frac{eg}{2m}\,\mathbf{S}\cdot\mathbf{B} = \frac{eg}{2m}\,[(S^2 - \xi^2)^{1/2}(B_x \sin\eta + B_y \cos\eta) + \xi B_z]. \quad \mathbf{(2.101)}$$

Then the equations of motion, Eq. 2.99, are equivalent to the canonical equations,

$$\frac{\partial H}{\partial \xi} = \dot{\eta}, \qquad \frac{\partial H}{\partial \eta} = -\dot{\xi}. \tag{2.102}$$

More generally, we can take the magnitude S also as one of the variables and introduce a variable ϕ conjugate to S. Because S is constant, we have the additional equations

$$\frac{\partial H}{\partial \phi} = \dot{S} = 0; \qquad \frac{\partial H}{\partial S} = \dot{\phi}. \tag{2.103}$$

The spin Lagrangian is given by

$$L = -H + \xi\dot{\eta} + S\dot{\phi} = -H + \xi\frac{\partial H}{\partial S} + S\frac{\partial H}{\partial S} = 0, \tag{2.104}$$

which is identically zero, but still allows one to derive the equations of motion.

The most satisfactory description of classical spin is in terms of a Lagrangian that simultaneously gives, both the equation of the motion for the trajectory, and for the spin. This cannot be achieved by a Lagrangian of the type Eq. 2.86 because it does not contain "coordinates" or "internal degrees of freedom" corresponding to spin whose variation would give the spin equation in the same way as the variation of the space time coordinates leads to the equation of the motion. We present now in this final paragraph a general method, independent of any specific interaction model, to obtain both the spin and momentum equations of the motion.[1]

Let us introduce as new internal coordinates four fourvectors

$$q^{\mu}_{(\alpha)} (\alpha = 0, 1, 2, 3)$$

with q_0 in the direction of the fourvelocity u_μ and the other three orthogonal to u_μ, hence space-like. Consider a Lagrangian that is a function of $\dot{x}$, q, $\dot{q}$ and the electromagnetic field:

$$L = L[\dot{x}_\mu, q^{\mu}_{(\alpha)}, \dot{q}^{\mu}_{(\alpha)}, A_\mu(x), A_{\mu,\nu}(x)]. \tag{2.105}$$

The variation of this Lagrangian leads to

$$\begin{aligned}
\delta L = {} & \left[\frac{\partial L}{\partial x^\mu} - \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right)\right] \delta x^\mu + \left[\frac{\partial L}{\partial q^\mu_\alpha} - \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}^\mu_{(\alpha)}}\right)\right] \delta q^\mu_{(\alpha)} \\
& + \frac{d}{ds}\left[\frac{\partial L}{\partial \dot{x}^\mu}\,\delta x^\mu + \frac{\partial L}{\partial \dot{x}^\mu}\frac{d}{ds}\,\delta x^\mu + \frac{\partial L}{\partial \dot{q}^\mu_{(\alpha)}}\,\delta q^\mu_{(\alpha)}\right] + \frac{\partial L}{\partial A_\lambda}(\delta A_\lambda - A_{\lambda,\mu}\,\delta x^\mu) \\
& + \frac{\partial L}{\partial A_{\lambda,\nu}}(\delta A_{\lambda,\nu} - A_{\lambda,\nu\mu}\,\delta x^\mu).
\end{aligned}$$

---

[1] I should like to thank James Currin for an enlightening discussion of this point.

The first two expressions in brackets are just the Euler-Lagrange equations for x and q respectively, and these will vanish.

If we now consider the translations

$$\delta x^{\mu} = a^{\mu}, \qquad \frac{d}{ds}\,\delta x^{\mu} = 0, \qquad \delta A_{\mu} = \delta q^{\mu}_{(\alpha)} = 0,$$

we obtain the equation of the motion

$$\dot{p}_{\mu} = \frac{\partial L}{\partial A_{\lambda}}\,A_{\lambda,\mu} + \frac{\partial L}{\partial A_{\lambda,\nu}}\,A_{\lambda,\nu\mu} \tag{2.106}$$

where

$$p_{\mu} = \partial L/\partial \dot{x}^{\mu}.$$

Next we consider the homogeneous Lorentz transformations

$$\delta x^{\mu} = \varepsilon^{\mu}{}_{\nu}x^{\nu}, \qquad \delta A_{\lambda} = \varepsilon_{\lambda}{}^{\nu}A_{\nu}, \qquad \delta A_{\lambda,\nu} = \varepsilon_{\lambda}^{\rho}A_{\rho,\nu} + \varepsilon_{\nu}{}^{\rho}A_{\lambda,\rho},$$

$$\delta q^{\mu}_{(\alpha)} = \varepsilon^{\mu}{}_{\nu}q^{\nu}_{(\alpha)},$$

and obtain[1]

$$\begin{aligned} &\frac{d}{ds}\left\{\frac{\partial L}{\partial \dot{q}^{[\mu}_{(\alpha)}}\,q^{(\alpha)}_{\nu]}\right\} + p_{[\mu}\dot{x}_{\nu]} \\ &\qquad + \left\{\frac{d}{ds}(p)_{[\mu}x_{\nu]} - \frac{\partial L}{\partial A_{\lambda}}\,A_{\lambda,[\mu}x_{\nu]} - \frac{\partial L}{\partial A_{\lambda,\rho}}\,A_{\lambda,\rho[\mu}x_{\nu]}\right\} \\ &\qquad = -\frac{\partial L}{\partial A_{[\mu}}\,A_{\nu]} - \frac{\partial L}{\partial A_{[\mu,\rho}}\,A^{\rho}_{\nu]} - \frac{\partial L}{\partial A_{\lambda,[\mu}}\,A_{\lambda,\nu]}. \end{aligned} \tag{2.107}$$

Using Eq. 2.106 we may rewrite Eq. 2.107 as follows

$$\dot{S}_{\mu\nu} + p_{[\mu}u_{\nu]} = -j_{[\mu}A_{\nu]} - M_{[\mu}^{\rho}A_{\nu],\rho} = M^{\lambda}{}_{[\mu}A_{\lambda,\nu]}$$

where

$$S_{\mu\nu} = \frac{\partial L}{\partial \dot{q}^{[\mu}_{(\alpha)}}\,q^{(\alpha)}_{\nu]}$$

$$j_{\mu} \equiv \partial L/\partial A^{\mu}$$

$$M^{\lambda\rho} = -M^{\rho\lambda} \equiv \partial L/\partial A_{\lambda,\rho}$$

---

[1] We use in the following a bracket around the subscripts for an antisymmetric tensor formed out of two vectors: $a_{[\mu}b_{\nu]} \equiv a_{\mu}b_{\nu} - a_{\nu}b_{\mu}$.

If the Lagrangian depends on the derivatives of A_μ only through the electromagnetic field tensor $F_{\mu\nu} = -F_{\nu\mu}$, $M_{\mu\nu}$ is antisymmetric and we can write

$$\dot{S}_{\mu\nu} + p_{[\mu}u_{\nu]} = -j_{[\mu}A_{\nu]} - M_{[\mu}{}^{\rho}F_{\nu]\rho}\,.$$

Now we assume as usual that the dependence of L upon A_μ is given by (See also Chapters III and IV)

$$\frac{\partial L}{\partial A^\mu} \equiv j_\mu = eu_\mu\,.$$

Defining a mechanical momentum G_μ by

$$G_\mu = p_\mu - eA_\mu$$

we obtain

$$\dot{G}_\mu = eF_{\lambda\mu}u^\lambda + \tfrac{1}{2}M^{\lambda\nu}F_{\lambda\nu,\mu}$$

$$\dot{S}_{\mu\nu} + G_{[\mu}u_{\nu]} = M_{\rho[\mu}F_{\nu]}{}^{\rho}$$

To solve for the mechanical moment we multiply the above equations by u^ν and obtain

$$G_\mu = m_0u_\mu + u^\nu S_{\mu\nu} - u^\nu M_{\rho[\mu}F_{\nu]}{}^{\rho},$$

where we have put

$$m_0 \equiv -u^\nu G_\nu\,,$$

then

$$\dot{S}_{\mu\nu} + u^\lambda \dot{S}_{[\mu\lambda}u_{\nu]} - u^\sigma M^\rho_{[\mu}F_{\sigma\rho}u_{\nu]}$$

$$+ u^\sigma M^\rho{}_\sigma F_{[\mu\rho}u_{\nu]} = M^\rho_{[\mu}F_{\nu]\rho}\,,$$

or finally

$$\dot{S}_{\mu\nu} + u^\lambda \dot{S}_{[\mu\lambda}u_{\nu]} = T_{\mu\nu} + u^\sigma T_{[\mu\sigma}u_{\nu]}\,, \qquad (2.108)$$

where

$$T_{\mu\nu} = M^\rho_{[\mu}F_{\nu]\rho}\,.$$

We notice that the previous equation is reduced to an identity upon multiplication by u_μ. This means that the six components of $S_{\mu\nu}$ are not all determined. However, because we assumed a Lagrangian independent of the accelerations $\dot{u}_\mu$ the so-called Weysenhoff condition

$$S_{\mu\nu}u^\nu = 0 \qquad (2.109)$$

is automatically satisfied. Now the three spatial components of spin are

3 1303 00076 3830

determined. We may further assume that the magnetic moment $M_{\mu\nu}$ is proportional to the spin

$$M_{\mu\nu} = mS_{\mu\nu}$$

One may now derive from Eq. 2.108, as a special case, the equation of Bargmann, Michel and Telegdi discussed above and the fact that $(m_0 + \sigma^{\mu\nu}F_{\mu\nu})$ is a constant of the motion.

5. *Further Developments and Exercises*

1. Consider relativistic energy conservation in nuclear reactions and verify the relations of Eq. 2.13 to Eq. 2.17.

2. Derive formulae connecting laboratory and center-of-mass frame energies and momenta, as well as scattering angles for two-body reactions.

3. By analogy to the photon define the neutrino as a mass-zero relativistic particle associated with the plane-wave solutions of the spinor-wave equation.

4. Solve the equation of motion, Eq. 2.22, of an electron moving in a fixed electromagnetic field in the simple case of constant electric and magnetic fields. (L. Landau and E. Lifshitz, *Classical Theory of Fields*, Addison-Wesley, Cambridge, Mass. (1951), p. 48–56.)

5. Solve the relativistic Kepler problem.

6. *A solution for a relativistic particle in an external scalar field.* Consider the Lagrangian, Eq. 2.45, or the equations of motion, Eq. 2.27. Assume that in a particular frame ϕ does not depend on x^0 (note that ϕ is a function of x^2); then

$$(m_0c^2 - \phi)\frac{dx^0}{ds} = \text{const.}$$

Choose $\phi = \phi(r) = g^2/r$ and the constant equal to m_0c^2:

$$\left(m_0c^2 - \frac{g^2}{r}\right) = \pm m_0c^2(1 - \beta^2)^{1/2}.$$

Show that if one chooses a + sign for $r > r_0$ and a − sign for $r < r_0$, one obtains a motion such that a particle approaches a force center until $r = r_0/2$, returns, and goes back to infinity. [J. Werle, *Bull. Acad. Pol.* 1, (1953) 281; *ibid.* 5, (1957) 491].

7. *Spinor Mechanics.* Introduce two component spinors ξ by the relation

$$u^\mu = \frac{dx^\mu}{ds} = \breve{\xi}_\alpha \sigma^{\mu\alpha\dot{\alpha}} \eta_{\dot{\alpha}},$$

where the right-hand side is a four-vector under proper Lorentz transformations. The free Lagrangian, with the Lagrange multipliers,

$$L_0 = \frac{1}{2}\left(\frac{d\xi_\alpha}{ds}\eta_{\dot{\alpha}} - \breve{\xi}_\alpha \frac{d\eta_{\dot{\alpha}}}{ds}\right) + \lambda_{pt}(\dot{x}^\sigma - \breve{\xi}\sigma^\mu \eta),$$

leads to the equations of motion,

$$\frac{d\breve{\xi}}{ds} = +\lambda_\mu \xi \sigma^\mu; \qquad \frac{d\eta}{ds} = -\hat{\lambda}_\mu \sigma^\mu \eta;$$

$$\frac{dx^\mu}{ds} = \xi\sigma^\mu \eta; \qquad \frac{d\hat{\lambda}^\mu}{ds} = 0.$$

If A is any 2×2 matrix, we get, with $[A, B] = AB - BA$,

$$\frac{d}{ds}(\xi A \eta) = \hat{\lambda}_\mu \breve{\xi}[\sigma^\mu, A]\eta.$$

If, for convenience, we wish to have parity conservation also, in the above equations we can replace ξ, η by four component spinors and σ^μ by the Dirac matrices γ^μ [A. Proca, *J. Phys. Radium* 15, (1954) 5], with $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$. The spin of the classical particle can be described by taking

$$A = \tfrac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu).$$

The interaction Lagrangian can be taken to be of the form

$$L_{\text{int}} = \xi O \eta,$$

where for electromagnetic interactions $O = eA_\mu\gamma^\mu$. Show that the Lorentz force can be derived from this equation.

8. *Thomas Precession.* Consider an electron moving in a circular orbit. In the equation of motion of a spin in a rest frame, Eq. 2.92, without the electric moment term,

$$\frac{d\mathbf{S}}{d\tau} = \left(\frac{g}{2}\right)\frac{e}{m}\mathbf{S} \times \mathbf{B},$$

the magnetic field refers to that in the rest frame. Let the electric and magnetic fields in the laboratory frame be $\mathbf{E}'$ and $\mathbf{B}'$ respectively. We must either transform $\mathbf{E}'$ and $\mathbf{B}'$ into a rotating frame of reference, or, if we transform them by a proper Lorentz transformation

$$\mathbf{B} \cong \left(\mathbf{B}' - \frac{\mathbf{v}}{c} \times \mathbf{E}'\right),$$

we must keep in mind that now we have a "rest frame" that is rotating. In the latter case, the equation of motion of the spin is

$$\frac{d\mathbf{S}}{d\tau} = \left(\frac{d\mathbf{S}}{d\tau}\right)_{\text{nonrotating}} - \boldsymbol{\omega}_T \times \mathbf{S},$$

where ω_T is the angular frequency of the Thomas precession, to be obtained as follows: Consider an electron in a circular path at two neighboring points P and P', with velocities $\mathbf{v}$ and $\mathbf{v} + \delta\mathbf{v}$ at these points. The laboratory frame is related to the rest frame at P by a Lorentz transformation with velocity $\mathbf{v}$, and to the rest frame at P' by a Lorentz transformation with velocity $\mathbf{v} + \delta\mathbf{v}$. The two rest frames at P and P' are related by a Lorentz transformation plus a rotation, and is of the form

$$\mathbf{x}'' = \mathbf{x}' + \mathbf{x}' \times \Delta\Omega;$$

therefore,

$$\boldsymbol{\omega}_T = \left[\frac{1}{(1-\beta^2)^{1/2}} - 1\right]\frac{\mathbf{v} \times \mathbf{a}}{v^2} \approx \frac{1}{2c^2}\,\mathbf{v} \times \mathbf{a}.$$

9. Discuss infinitesimal canonical transformations in covariant form for a single particle and the relation of their generators to the constants of motion.

10. Write the covariant Hamilton-Jacoby equation for a single particle.

11. Derive Eqs. 2.102, 2.103, and 2.104.

12. Consider a system of particles, and let $f(\mathbf{p})d^3p$ be the number of particles with momentum in the range $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$. Show that $f(\mathbf{p})E$ is an invariant quantity under Lorentz transformations. Here E is the energy of the particle. *Hint*: Use the invariance of d^4p and the relation

$$d^4p\,\delta(p^2 - m^2) = d^3p(m/E)$$

13. Show that for two colliding monochromatic beams of particles the invariant current density is given by

$$j = [(k_1 \cdot k_2)^2 - m_1^2 m_2^2]^{1/2}/m_1 m_2$$

where k_1 and k_2 are the four-momenta of the particles and m_1 and m_2 their masses. *Hint*: Take a special frame in which the particles are moving towards each other along a straight line. In this frame the current density is $j = (v_1 + v_2)/V_1 V_2$, where v_i $(i = 1, 2)$ are the velocities of the particles, $v_i = p_i/E_i$ (p is the magnitude of the three-momentum, E the energy) and $V_i = m_i/E_i$ are the invariant normalization volume.

14. Extend problems 12 and 13 to the case of mass zero particles.

15. Let P be a point on the world line of a particle. Introduce at this point an orthogonal basis (Vierbein or fourleg)

$$g^{\mu\nu} e_\mu^{(\alpha)} e_\nu^{(\beta)} = g^{\alpha\beta}$$

$$e_\mu^{(0)} = u_\mu$$

where u_μ is the fourvelocity of the particle. Show that under the equation of motion

$$\frac{d}{ds}\, e_\mu^{(i)} = u_\mu e_\nu^{(i)} \frac{du^\nu}{ds}$$

$$e^{(0)}_\mu(s) = u_\mu(s)$$

the basis remains an orthogonal basis along the world line of the particle.

16. The so-called *Fermi Coordinates* of a point A with respect to a world line with the velocity vector $u_\mu(s)$ are defined as follows:

$$x^{(i)} = dn^\mu e_\mu^{(i)}$$

$$x^{(0)} = s$$

where d is the distance of A from the worldline (distance of the geodesic in the case of general relativity) and $e_\mu^{(\alpha)}$ is the basis introduced in the previous problem, n^μ the normal to the world line at the intersection of the normal to the point A, and s is the length of the world line from some origin P to the intersection point. Show that these coordinates are invariant.

BIBLIOGRAPHY FOR CHAPTER II

In addition to books listed at the end of Chapter I, the following articles may be consulted:

1. H. C. Corben, "Spin in Classical and Quantum Theory," *Phys. Rev., 121* (1961), 1833.
2. J. Frenkel, "Zur Elektrodynamik Punktförmiger Elektronen," *Z. Physik, 32* (1925), 518.
3. J. Frenkel, "Die Elektrodynamik des Rotierenden Elektrons," *Z. Physik, 37* (1926), 243.
4. F. Halbwachs, "Lagrangian Formalism for a Classical Relativistic Particle Endowed with Internal Structure," *Prog. Theor. Phys., 24* (1960), 291.
5. G. Kallman, "Lagrangian Formalism in Relativistic Dynamics," *Phys. Rev., 123* (1961), 384.
6. P. M. Quam, "Sur le Dynamique Analytic du Point en Relativité," *Acta Phys. Austriaca, 14* (1961), 232.

7. A. Peres and N. Rosen, "Covariant Formalism for Particle Dynamics," *Nuovo Cimento, 18* (1960), 644.
8. A. Proca, "Mécanique du Point," *J. Phys. Radium, 15* (1954), 65.
9. F. Sauter, "Zur Lorentz-Invarianten Formulierung der Kanonischen Bewegungsgleichungen in der Punktmechanik," *Z. Physik, 156* (1959), 275.
10. R. Schiller, "Quasi-classical Theory of Relativistic Spinning Electron," *Phys. Rev., 128*, 1402 (1962) and *125*, 1116 (1962).
11. V. L. Telegdi and R. Winston, "A Dynamical Interpretation of the Thomas Precession," *Beiträge zur Entwicklung der Physik* (Basel: Birkhäuser Verlag, 1960).

III.

RELATIVISTIC FIELD THEORY

1. *Intuitive Introduction of Fields*

We have already met the concept of *field* in a limited sense, namely as an external-force field on a particle. We have spoken of given external scalar or electromagnetic fields, producing an "action-at-a-distance" type of force on the particle. More fundamentally, a field is a mechanical system in its own right, carries energy and has, in general, a continuously infinite number of degrees of freedom. It is a system filling the whole space like a "fluid", and it has physical reality. Whereas in the previous chapter the mass point was the physical system under consideration and the fields were merely auxiliary (phenomenological) quantities, in this chapter the fields will be our physical system, and we shall describe and seek the laws obeyed by the fields themselves. The more realistic (classical) description of the physical systems must consider the interacting fields and particles together. This situation will be discussed in the second part of the book. However, if the interaction with the particles is weak, the laws of the so-called *free fields* can be observed directly, as in the case of a radiation field, for example.

The infinite number of degrees of freedom of the field must be described by continuous indices. Instead of the coordinates $q_1, q_2, \cdots$, the dynamical variables of the field will be a set of functions $\psi^\alpha(\mathbf{x}, t)$, $\alpha = 1, \cdots, N$, where $(\mathbf{x}, t)$ are now parameters which, together with α, label the degrees of freedom of the system. It is then possible to formulate a completely relativistic and elegant theory. However, some divergence difficulties are present with a continuously infinite number of degrees of freedom. Although it is rather straightforward to formulate a relativistic theory for a single particle or, at the other extreme, for a system with an infinite number of degrees of freedom, it seems very difficult to formulate a relativistic theory for a finite number of particles.

We first illustrate the passage from a mechanical system to a field in terms of a simple one-dimensional model. Consider a linear chain of mass points, and let ψ_i be the displacement of the i^{th} particle from its equilibrium position, and let a be the equilibrium separation between the mass points. Let us assume that forces acting on each particle come from the two neighboring particles and that the effect of the remaining mass points is negligible. For simplicity we consider the force between two mass points to be proportional to their separations. The force on the i^{th} particle is, thus,

$$F_i = -\frac{\partial V}{\partial \psi_i} = k(\psi_{i+1} - \psi_i) - k(\psi_i - \psi_{i-1}).$$

The Lagrangian of the total system is given by

$$L = \tfrac{1}{2} \sum_i [m_i \dot{\psi}_i^2 - k(\psi_{i+1} - \psi_i)^2],$$

where the dots on the dynamical variables ψ_i denote time derivatives. We have a discrete but infinite number of degrees of freedom, and our treatment at present is nonrelativistic. For each mass point i the above Lagrangian yields the equation of motion

$$m_i \ddot{\psi}_i - k(\psi_{i+1} - \psi_i) + k(\psi_i - \psi_{i-1}) = 0.$$

Now we go to the limit $a \to 0$ in such a way that m_i/a approaches a finite value $\rho(x)$. The discrete index i becomes a continuous index x. Then we have the following limiting relations:

$$\psi_i(t) \to \psi(x, t),$$

$$\frac{\psi_{i+1} - \psi_i}{a} \to \frac{\partial \psi}{\partial x},$$

$$\frac{1}{a}\left(\frac{\psi_{i+1} - \psi_i}{a} - \frac{\psi_i - \psi_{i-1}}{a}\right) \to \frac{\partial^2 \psi}{\partial x^2},$$

$$ka \to Y,$$

where Y is a constant, namely Young's modulus, because it is the proportionality constant between the force and the displacement per unit length:

$$F = ka \frac{\psi_{i+1} - \psi_i}{a}.$$

The limit of the Lagrangian is, then,

$$L = \tfrac{1}{2}\int dx\left[\rho\dot{\psi}^2 - Y\left(\frac{\partial\psi}{\partial x}\right)^2\right] = \int dx\,\mathscr{L}(x, t),$$

where we have introduced a Lagrangian *density* $\mathscr{L}(x, t)$. In this limit the action principle $\delta\int L\,dt = 0$ becomes

$$\delta\int \mathscr{L}\,dx\,dt = 0.$$

Instead of the variation of the coordinates ψ_i, we have now a variation of the function $\psi(x, t)$. In mathematical language we have here a variational problem in two variables (x and t). The mathematical theory of the variational problem in several variables and for several functions is standard and very straightforward.[1]

The limit of the equation of motion is

$$\rho\ddot{\psi} - Y\frac{\partial^2\psi}{\partial x^2} = 0,$$

which is a wave equation with a propagation velocity $v = \sqrt{Y/\rho}$. Here $\psi(x, t)$ represents the amplitude of the elastic waves. The wave equation coincides with the Euler-Lagrange equations,

$$\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial\dot{\psi}}\right) - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial\psi^{,x}}\right) - \frac{\partial L}{\partial\psi} = 0,$$

where $\psi^{,x} = \partial\psi/\partial x$, which itself is the limit of the Euler-Lagrange equations for the discrete system.

Generalizing this example, we see that we can define the dynamics of the field by an action principle

$$\delta\int \mathscr{L}\,d^3x\,dt = 0,$$

where the Lagrangian density $\mathscr{L}$ is a function of the field variables $\psi^\alpha(x, t)$, $\alpha = 1, \cdots, N$, and their derivatives. It is clear that if $\mathscr{L}$ is a Lorentz scalar, the action principle will take an invariant form

$$\delta\int \mathscr{L}\,d^4x = 0, \tag{3.1}$$

and the resultant field equations will be invariant under Lorentz transformations. It is this fact that makes possible a relativistic theory of a system with continuously infinite degrees of freedom.

Before we elaborate on the action principle, Eq. 3.1, we introduce in the next section the prototype and the most important of all fields, namely the electromagnetic field. We discuss inductively the basic equations and demonstrate their relativistic invariance.

---

[1] See, for example, R. Courant and D. Hilbert, *Methods of Mathematical Physics* (New York: Interscience Publishers, 1953), Vol. I.

2. *The Electromagnetic Field*

The description of the interaction between charged particles by the intermediary of a field as an energy-carrying physical system goes back to Faraday and Maxwell. The field is produced by the charged particles—whose existence we assume as the starting point—and is measured by the acceleration that it produces when acting (again) on charged particles. The concept of the electromagnetic field is perhaps best introduced *deductively* by its basic equations, Maxwell's equations and the auxiliary equations. However, we shall proceed in a slightly inductive manner.

Basic Equations[1]

Charged particles at rest are acted on by a force $e\mathbf{E}$, where e is an attribute of the particle, its charge.[2] Therefore the electric field $\mathbf{E}$ may be introduced as the electric force per unit charge in the same way as the gravitational field $\mathbf{g}$ is defined to be the gravitational force per unit mass. Similarly, the magnetic field $\mathbf{B}$ may be formally introduced as the force per unit magnetic pole. However, the meaning of the magnetic monopoles is not very clear, so we go back to the general expression for the force acting on a moving particle discussed in Chapter II. If the charged particle is in motion, the force acting on it is given by

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{3.2}$$

This equation may be taken as the definition of the electric and magnetic fields. We also know from the conservation of energy that $\mathbf{E} = -\nabla\phi$, or $\nabla \times \mathbf{E} = 0$, and that the magnetic field $\mathbf{B}$ does not do any work. We emphasized in Chapter II that $\mathbf{E}$ and $\mathbf{B}$ are different aspects of the same thing, the electromagnetic field $F_{\mu\nu}$.

In material media the existence of many charged particles, and of atomic and nuclear structures, makes a continuous description difficult. This model of matter is a quantum-mechanical one. Nevertheless, one can introduce, phenomenologically, two new quantities $\mathbf{D}$ defined by

$$\oint_F \mathbf{D}\cdot d\mathbf{f} = Q, \tag{3.3}$$

---

[1] In this section we use rationalized mks units, the four fundamental units length l (meter), mass m (kilogram), time t (second) and charge q (coulomb) and $\varepsilon_0\mu_0 = 1/c^2$. Of course, the number of fundamental units or the choice of them is arbitrary and has no fundamental significance. (Later on we shall find it convenient to put $c = 1$.)

[2] For the interpretation of e as the coupling constant between the charged particle and the electromagnetic field, see Part II.

where Q is the total charge inside the closed surface F, and $\mathbf{H}$ = pole strength/unit area. Presently we shall give alternate expressions and interpretations for these quantities.

The fundamental postulates of electrodynamics in material media (Maxwell's electrodynamics) are:

(1) Faraday's law,

$$\frac{d}{dt}\int_F \mathbf{B}\cdot\mathbf{df} = -\oint_C \mathbf{E}\cdot\mathbf{ds}\,; \tag{3.4}$$

(2) Ampère's law,

$$\int_F \mathbf{C}\cdot\mathbf{df} = \oint_C \mathbf{H}\cdot\mathbf{ds}, \tag{3.5}$$

where $\mathbf{C} = \mathbf{J} + d\mathbf{D}/dt$, and $\mathbf{J}$ is the current vector.

Faraday's law is a kind of relativity effect with respect to the motion. If we move a closed circuit in a magnetic field with a velocity $\mathbf{v}$, the charges will be acted on by a force $\mathbf{v} \times \mathbf{B}$. On the other hand, if we hold the circuit fixed and move the "magnets," the same effect should occur. But now, because the charges are at rest, they must be acted on by an electric force $e\mathbf{E}$. The equality of these two effects is expressed by Eq. 3.4.[1]

In particular, for closed surfaces Eqs. 3.4 and 3.5 become

$$\oint \mathbf{B}\cdot\mathbf{df} = \text{constant in time};$$

$$\oint \mathbf{C}\cdot\mathbf{df} = 0. \tag{3.6}$$

The second of these equations indicates that the electric current lines are closed. Because B is finite and the space is homogeneous, the same must hold for magnetic lines in agreement with observation; i.e.,

$$\oint \mathbf{B}\cdot\mathbf{df} = 0. \tag{3.7}$$

From Eqs. 3.6 and 3.3, we obtain

$$\frac{d}{dt}\oint \mathbf{D}\cdot\mathbf{df} + \oint \mathbf{J}\cdot\mathbf{df} = 0,$$

or

$$\frac{dQ}{dt} + \oint \mathbf{J}\cdot\mathbf{df} = 0. \tag{3.8}$$

This is the *equation of continuity*, which implies that if $\mathbf{J} = 0$, the total charge Q is constant. One can furthermore introduce a charge density ρ defined by

$$\int_F \mathbf{D}\cdot\mathbf{df} = \int_V \rho\, dV, \tag{3.9}$$

---

[1] See Exercise 1.

or, by Gauss's theorem,

$$\rho = \nabla \cdot \mathbf{D}; \tag{3.9'}$$

then the equation of continuity, Eq. 3.8, can be written in differential form as

$$\dot{\rho} + \nabla \cdot \mathbf{J} = 0. \tag{3.10}$$

The three basic equations, Eqs. 3.4, 3.5, and 3.7 can also be written in differential form:

$$\begin{gathered} \dot{\mathbf{B}} = -\nabla \times \mathbf{E}; \\ \dot{\mathbf{C}} = \dot{\mathbf{D}} + \mathbf{J} = \nabla \times \mathbf{H}; \\ \nabla \cdot \mathbf{B} = 0. \end{gathered} \tag{3.11}$$

However, these seven equations are not sufficient to determine the electromagnetic field (**E, B, D, H**) for a given **J**. One has then to make use of empirical or phenomenological and approximate relations.

Empirically, charges in free space and at rest produce an electric field given by

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho}{r^3} \mathbf{r}\, d^3x, \tag{3.12}$$

from which we obtain

$$\nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0} \rho. \tag{3.13}$$

Similarly, stationary currents produce a magnetic field given by

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int (\mathbf{J} \times \mathbf{r}/r^3)\, d^3x, \tag{3.14}$$

which leads to the relation

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \tag{3.15}$$

True to their empirical character, Eqs. 3.13 and 3.15 contain two new empirical constants ε_0 and μ_0. So far we have used charges and currents in free space. If matter is present, charges and currents that make up matter will be influenced by external electromagnetic fields. The charges will be displaced from their equilibrium position, giving rise to electric moments and polarization. The *polarization* **P** is defined as the electric moment per unit volume. The displaced charges will produce an additional electric field whose effective density is given by

$$\frac{1}{\varepsilon_0} \rho' = -\nabla \cdot \mathbf{P}. \tag{3.16}$$

Therefore, in matter, instead of Eq. 3.13, we will have the relation[1]

$$\nabla\cdot\mathbf{E} = \frac{1}{\varepsilon_0}(\rho + \rho'),$$

or

$$\rho = \nabla\cdot(\varepsilon_0\mathbf{E} + \mathbf{P}) = \nabla\cdot(\varepsilon\mathbf{E}). \tag{3.17}$$

If we compare Eq. 3.17 with Eq. 3.9′, we get the first of our empirical relations

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P} \equiv \varepsilon\mathbf{E}. \tag{3.18}$$

This equation defines the *dielectric constant* of matter in terms of ε_0 and **P**.

Similarly, the circular orbits of the charges inside matter constitute a magnetic moment **M**, which can be related to an effective new current by

$$\mathbf{J}' = \nabla \times \mathbf{M}. \tag{3.19}$$

Therefore, in matter, instead of Eq. 3.15, we have the relation

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \mathbf{J}'),$$

or

$$\mathbf{J} = \nabla \times \left(\frac{1}{\mu_0}\mathbf{B} - \mathbf{M}\right). \tag{3.20}$$

We define a new quantity

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M} \equiv \frac{1}{\mu}\mathbf{B}, \tag{3.21}$$

which we introduced before without much elaborating. Here μ is the *magnetic permeability* of matter. We have, then,

$$\nabla \times \mathbf{H} = \mathbf{J};$$

if, in addition to the stationary currents **J**, we have displacement currents given by $\dot{\mathbf{D}}$, we obtain the second of the fundamental equations, Eq. 3.11, or Ampère's law. At the same time, Eq. 3.21 is the second of our empirical relations.

It is important to note that an essentially quantum-mechanical model of matter, with its bound and circulating charges, was used to derive the above empirical relations. A mixture of positive or negative charges, as in a plasma, would behave quite differently. Any further investigation of **D** and **H**, and hence ε and μ, must make use of the detailed structure of matter.

Historically, **H** played the role of the magnetic field vector instead of **B**. The reason for this is as follows. The mechanical momentum in an external magnetic field **B**, due to the magnetic moment **M** of matter, is

---

[1] See Exercise 2.

$$\mathbf{B} \times \mathbf{M} = (\mu_0\mathbf{H} + \mathbf{M}) \times \mathbf{M} = \mu_0\mathbf{H} \times \mathbf{M} = \mathbf{H} \times \mathbf{M}',$$

where we have introduced $\mathbf{M}' = \mu_0\mathbf{M}$. Thus the magnetic moment seems to be evaluated with **H** if we use **M**′ as the effective magnetic moment.

Usually a third empirical relation is introduced, namely Ohm's law,

$$\mathbf{J} = \sigma\mathbf{E}. \tag{3.22}$$

This relation is on a slightly different basis than the previous two empirical equations are, because we may solve Eqs. 3.11 with **J** given, without using Ohm's law.

With the phenomenological equations, it is then possible to solve the fundamental equations, Eq. 3.11. For example, the elimination of **D** and **H** gives

$$\left(\varepsilon\mu \frac{\partial^2}{\partial t^2} + \mu\sigma \frac{\partial}{\partial t}\right)\mathbf{E} = -\nabla \times (\nabla \times \mathbf{E}),$$

in the special case $\rho = 0$, $\nabla\cdot\mathbf{E} = 0$,

$$\varepsilon\mu \frac{\partial^2\mathbf{E}}{\partial t^2} + \sigma\mu \frac{\partial\mathbf{E}}{\partial t} = \Delta E\,;$$

exactly the same equations hold for **B**.

Lorentz Electrodynamics

We shall consider now the so-called "microscopic equations" for point charges in vacuum. No complications due to the atomic structure of matter are present. We assume that the same basic laws are valid microscopically as was the case macroscopically. Because no matter is present, we can identify **D** and **E**, **B** and **H**. The current **J** is due entirely to the motion of charged particles; hence $\mathbf{J} = \rho\mathbf{v}(\mathbf{x}t)$, where $\mathbf{v}(\mathbf{x}t)$ is the velocity of the charges at point **x** at time t. The fundamental equations become

$$\begin{aligned} \dot{\mathbf{B}} &= -\nabla \times \mathbf{E}\,; \\ \dot{\mathbf{E}} + \rho\mathbf{v} &= \nabla \times \mathbf{B}\,; \\ \nabla\cdot\mathbf{B} &= 0\,; \\ \nabla\cdot\mathbf{E} &= \rho, \end{aligned} \tag{3.23}$$

where the two functions, velocity and charge distribution are related by the *equation of continuity*

$$\dot{\rho} + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{3.24}$$

Actually these microscopic equations must be taken as the more fundamental equations. The macroscopic equations, Eqs. 3.11, must be obtained from here by some statistical averaging process, using the structure of material media. This is indeed the case.[1]

In particular, we see from Eq. 3.23 that for empty space $\rho = 0$, **E** and **B** satisfy the *wave equation*

$$\Box^2\mathbf{E} = 0, \qquad \Box^2\mathbf{B} = 0. \tag{3.25}$$

The system of equations, Eqs. 3.23, can be further simplified by the introduction of the *potentials* ϕ and **A** defined by

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}, \qquad \mathbf{B} = \nabla \times \mathbf{A}, \tag{3.26}$$

which already have been discussed briefly in Chapter II. The first and third equations in Eqs. 3.23 are satisfied automatically, and we are left with the following two equations:

$$\begin{aligned} \Delta\mathbf{A} - \nabla(\nabla\cdot\mathbf{A}) &= -\frac{1}{c}\frac{\partial}{\partial t}\left(\nabla\phi + \frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}\right) + \frac{\rho}{c}\mathbf{v}; \\ \nabla\cdot\left(\nabla\phi + \frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}\right) &= -\rho. \end{aligned} \tag{3.27}$$

However, the simplification obtained by the introduction of ϕ and **A** must be paid for by the fact that ϕ and **A** are not unique for a given field (**E**, **B**). We shall discuss this point in detail later.

Next, we shall discuss the covariant form of the fundamental equations in order to demonstrate the relativistic invariance of electrodynamics.

Covariant Form of Maxwell-Lorentz Equations

From the potentials ϕ, **A**, we pass to the four-vector potentials

$$A^\mu : (\phi, \mathbf{A}); \tag{3.28}$$

then the defining equations, Eqs. 3.26, can be written in simple tensor form:

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}, \tag{3.29}$$

as has been discussed already in Chapter II, Eqs. 2.20 and 2.21.

In addition, we make the physical postulate that the total charge Q, or equivalently the equation of continuity, is invariant under Lorentz transformations. Evidently the value of the total charge Q is the same in every coordinate frame. In analogy to the four-momentum of a particle $p^\mu = m_0 c u^\mu$, we define now a *four-current density* j^μ by

---

[1] See Exercise 7.

$$j^\mu \equiv \rho_0 c u^\mu \equiv \rho \frac{dx^\mu}{dt}, \tag{3.30}$$

where ρ_0 is a rest-charge density and is an invariant quantity, and

$$\rho = \rho_0 c \frac{dt}{ds}$$

is the actual charge density which transforms like energy (zero component of a four-vector). The equation of continuity can now be written as

$$j^\mu_{,\mu} = \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{3.31}$$

and is simply the four-dimensional divergence of the current vector. We also note that the current vector always has a constant length:

$$j^\mu j_\mu \equiv j^2 = \rho_0^2 c^2. \tag{3.32}$$

In terms of the $F_{\mu\nu}$ and the j^μ the first and the third equations of Eqs. 3.23 are indeed satisfied automatically, for any antisymmetric tensor $F_{\mu\nu}$ satisfies the identity

$$\varepsilon^{\mu\nu\sigma\rho} F_{\nu\sigma,\rho} \equiv [F_{\nu\sigma,\rho}]_{[\nu\sigma\rho]} = 0,$$

or

$$F_{\nu\sigma,\rho} + F_{\sigma\rho,\nu} + F_{\rho\nu,\sigma} = 0. \tag{3.33}$$

The remaining two equations, Eq. 3.27, can be written in the simple form

$$F^{\mu\nu}{}_{,\nu} = -\frac{1}{c} j^\mu, \tag{3.34}$$

which is the covariant form of the Maxwell equations. The last equation, by the way, also contains the equation of continuity, for $F^{\mu\nu}{}_{,\nu\mu}$ is automatically zero because $F^{\mu\nu}$ is antisymmetric.

Gauge Transformations

We have already mentioned that potentials do not determine the electromagnetic field uniquely. A transformation of potentials of the form

$$A'_\mu = A_\mu + \frac{\partial \Lambda}{\partial x^\mu}, \tag{3.35}$$

where $\Lambda(x)$ is an arbitrary scalar function, does not effect the field components. The field calculated from A'_{μ} is the same as the one calculated from A_{μ}:

$$F'_{\mu\nu} = A_{\nu,\mu} + \Lambda_{,\nu\mu} - A_{\mu,\nu} - \Lambda_{,\mu\nu} = F_{\mu\nu}.$$

Therefore, only those quantities invariant under the *gauge transformations*, Eq. 3.35, will have direct experimental significance. Such quantities or expressions will be called *gauge invariant.*

On the other hand, we can make use of the freedom given by gauge transformations to simplify the field equations, Eqs. 3.34. If we choose A_{μ} such that

$$\frac{\partial A^{\mu}}{\partial x^{\mu}} = 0, \qquad (\textit{Lorentz gauge}) \tag{3.36}$$

then the field equations become simply

$$\Box^2 A^{\mu} = \frac{1}{c}\, j^{\mu}. \tag{3.37}$$

The choice of Lorentz gauge can always be made by adding to A_{μ} a term $\Lambda_{,\mu}$ such that $\Lambda_{,\mu}{}^{\mu} = -A_{\mu}{}^{,\mu}$. Even with the choice of Lorentz gauge, the field is not determined uniquely: we can add to the potential A^{μ} the divergence $f_{,\mu}$ of any solution of the d'Alembert equation $\Box^2 f = 0$, without changing the field.

In the second part of this book, we shall also use a different gauge:

$$\nabla \cdot \mathbf{A} = 0, \qquad (\textit{radiation gauge}). \tag{3.36}$$

Another useful choice of potential is

$$\phi = 0, \qquad \nabla \cdot \mathbf{A} = 0, \qquad (\textit{Coulomb gauge}). \tag{3.36$'$}$$

Invariants

From the vector A^{μ} or tensor $F^{\mu\nu}$ and their derivatives, many quantities invariant under Lorentz transformations can be constructed. Two important invariants containing only the fields $F^{\mu\nu}$, which therefore are also gauge invariant, are

$$\begin{gathered} F^{\mu\nu}F_{\mu\nu}, \\ \varepsilon^{\mu\nu\lambda\rho}F_{\mu\nu}F_{\lambda\rho}, \end{gathered} \tag{3.38}$$

or, in terms of the electric and magnetic fields, $(E^2 - B^2)$ and $\mathbf{E}\cdot\mathbf{B}$. As a consequence, we have statements of the form that if $\mathbf{E}$ and $\mathbf{B}$ are orthogonal or have equal magnitudes in one frame of reference, they are orthogonal and have equal magnitudes in all frame of references.

Covariant Form of the Maxwell Equations in Material Media

If we introduce two new antisymmetric tensors[1]

$$G^{\mu\nu} = \begin{pmatrix} 0 & -cD_1 & -cD_2 & -cD_3 \\ & 0 & -H_3 & H_2 \\ & & 0 & -H_1 \\ & & & 0 \end{pmatrix}, \qquad K^{\mu\nu} = \begin{pmatrix} 0 & cB_1 & cB_2 & cB_3 \\ & 0 & -E_3 & E_2 \\ & & 0 & -E_1 \\ & & & 0 \end{pmatrix}, \tag{3.39}$$

and a current density four-vector

$$j^\mu = (c\rho, \mathbf{J}), \tag{3.40}$$

then the fundamental Eqs. 3.11, together with $\nabla \cdot \mathbf{D} = \rho$, can be written in the form

$$\begin{aligned} G^{\mu\nu}{}_{,\nu} &= -j^\mu, \\ K^{\mu\nu}{}_{,\nu} &= 0. \end{aligned} \tag{3.41}$$

The two tensors are coupled by the phenomenological Eqs. 3.18 and 3.21, or

$$G^{\mu\nu} = E^{\mu\nu\sigma\rho} K_{\sigma\rho}, \tag{3.42}$$

where, in the case of homogeneous media, $E^{\mu\nu\sigma\rho}$ is the complete antisymmetric symbol with elements

$$\begin{aligned} E^{0\nu\sigma\rho} &= c\varepsilon\,; \qquad \nu = 1, 2, 3\,; \\ E^{kl\sigma\rho} &= 1/\mu c\,; \quad k = 1, 2, 3. \end{aligned}$$

If the material medium is not homogeneous, $E^{\mu\nu\sigma\rho}$ is a more complicated tensor, ε and μ being 3 × 3 tensors themselves. From the transformation property of $E^{\mu\nu\sigma\rho}$, we can thus obtain the transformation properties of ε_{ij} and μ_{ij} under Lorentz transformations.[2] The phenomenological Ohm's law gives another connection between $G^{\mu\nu}$ and $K_{\mu\nu}$, namely,

$$G^{k\mu}{}_{,\mu} = \sum{}^{klm} K_{lm}, \tag{3.43}$$

where $\sum^{klm}$ is, for homogeneous media, the completely antisymmetric tensor, with $\sum^{123} = \sigma$.

Again we can introduce potentials

$$\mathfrak{a}^\mu : (\mathfrak{a}^0, \mathfrak{a})\,; \qquad \mathfrak{b}^\mu = (\mathfrak{b}^0, \mathfrak{b}),$$

such that

$$G_{\mu\nu} = \mathfrak{a}_{\nu,\mu} - \mathfrak{a}_{\mu,\nu}\,; \qquad K_{\mu\nu} = \mathfrak{b}_{\nu,\mu} - \mathfrak{b}_{\mu,\nu},$$

---

[1] For the corresponding forms of $G_{\mu\nu}$ and $G^{\mu}{}_{\nu}$, ···, see Eq. 1.104.

[2] Exercise 6.

in order to simplify the field equations to the form

$$\begin{aligned}\Box^2 \mathfrak{a}^\mu &= j^\mu,\\ \Box^2 \mathfrak{b}^\mu &= 0,\end{aligned} \tag{3.44}$$

by a suitable choice of potential.

Spinor Form of the Maxwell Equations*[1]

The covariant form of the Maxwell equations in terms of the vector field $A_\mu(x)$ is the most natural one; however, it is not the only possible covariant form. We have seen in Chapter I that there are other ways to express the Lorentz invariance of a theory. In particular, it was emphasized that the spinor representations of the Lorentz group, in terms of two-component spinors, were more fundamental in the sense that vector and tensor representations can be built up by combinations of these spinors. Therefore it is interesting, and for some purposes useful, to have a spinor representation of the electromagnetic field.

There are various ways of writing the Maxwell equations in spinor form. The most natural way is to use the correspondence between tensors and spinors discussed in Chapter I. According to this correspondence [see Eq. 1.60, Chapter I], a vector index is equivalent to two spinor indices, one dotted and the other undotted. For the electromagnetic field we have, therefore,

$$F^{\mu\nu} \leftrightarrow F^{\alpha\dot{\alpha}\beta\dot{\beta}},$$

where

$$F^{\alpha\dot{\alpha}\beta\dot{\beta}} = \sigma_\mu^{\alpha\dot{\alpha}}\ \sigma_\nu^{\beta\dot{\beta}} F^{\mu\nu}\,.$$

In addition, we know that $F^{\mu\nu}$ is a real antisymmetric tensor.

This condition allows us to describe the electromagnetic field by a second-rank spinor, rather than fourth. For, if $F^{\mu\nu}$ is real and skew symmetric, there exists a uniquely determined *symmetric* second-rank spinor $\phi^{\alpha\beta}$ such that

$$F^{\alpha\dot{\alpha}\beta\dot{\beta}} = \tfrac{1}{2}[\phi^{\alpha\beta} C^{\dot{\alpha}\dot{\beta}} + C^{\alpha\beta}\bar{\phi}^{\dot{\alpha}\dot{\beta}}].$$

Note that on the right-hand side C is antisymmetric, so that F is split into a part symmetric in (α, β), antisymmetric in $(\dot{\alpha}, \dot{\beta})$ and a part symmetric in $(\dot{\alpha}, \dot{\beta})$ and antisymmetric in (α, β).

Because $\phi_{\alpha\beta}$ is uniquely determined, the electromagnetic field can be described by a symmetric second-rank spinor. If we define derivatives with respect to spinor indices by the correspondence

$$\partial^{\alpha\dot{\alpha}} = \sigma_\mu^{\alpha\dot{\alpha}}\partial^\mu,$$

---

[1] Sections marked with an asterisk may be omitted at first reading.

and the equivalent of the current vector by

$$^{\alpha\dot{\alpha}} = \sigma_\mu^{\alpha\dot{\alpha}} j^\mu,$$

the Maxwell equations, Eqs. 3.34 can be written in purely spinorial form:

$$\partial^{\alpha\dot{\alpha}} \phi_{\alpha\beta} = -\frac{1}{c} j_\beta^{\dot{\alpha}}. \tag{3.45}$$

The Maxwell equations can also be rewritten in the form of neutrino equations. If we define a four-vector G_μ by

$$\mathbf{G} = \mathbf{E} + i\mathbf{B}, \qquad \text{and} \qquad G_0 = 0,$$

then we can write the Maxwell equations in the form

$$\beta_\nu \frac{\partial G_\mu}{\partial x_\nu} = -\frac{1}{c} j_\mu, \tag{3.46}$$

where β_μ are four 4 × 4 hermitian matrices obeying the relations

$$\beta_0 \beta_\mu = \beta_\mu \beta_0, \qquad \beta_\mu^2 = 1; \qquad \mu = 0, 1, 2, 3;$$

$$\beta_i \beta_j = \delta_{ij} - i\varepsilon_{ijk}\beta_k; \qquad i, j, k = 1, 2, 3; \varepsilon_{123} = 1.$$

The explicit form of the matrices are given by

$$\beta_0 = I \times I; \qquad \beta_1 = \begin{pmatrix} -\sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix} = -\sigma_3 \otimes \sigma_2;$$

$$\beta_2 = i\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = +\sigma_2 \otimes I; \qquad \beta_3 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix} = \sigma_1 \otimes \sigma_2,$$

where the σ's are Pauli matrices. Thus the β matrices are the Kroenecher products (⊗) of the Pauli matrices.

By a similarity transformation on β's of the form $\beta'_\mu = S^{-1}\beta_\mu S$, with

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -i & 0 & i & 0 \\ 0 & i & 0 & -i \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

we can reduce the β matrices fully to

$$\beta'_0 = \beta_0; \qquad \beta'_1 = \sigma_3 \otimes \sigma_3; \qquad \beta'_2 = -\sigma_3 \otimes \sigma_1; \qquad \beta'_3 = -I_1 \otimes \sigma_2,$$

so that the four-component Eq. 3.46 splits into two two-component equations (see Exercise 11). Unlike the neutrino equation, these two equations are coupled through the subsidiary condition $G_0 = 0$, which we introduced above.

Equation 3.46 has the general form of the wave equations for all spin values. The Dirac equation and various forms of the so-called Kemmer equations have the same form.

3. *Lagrangian Form of Field Equations*

We have seen that fields as physical systems will be described by a set of functions $\psi^\alpha(\mathrm{x}, t)$, $\alpha = 1, \cdots, N$, satisfying certain partial differential equations, the so-called field equations. (Instead of the initial conditions in ordinary differential equations of mechanics, we have here boundary conditions to ensure a unique solution.) For example, the field in the example given in Sect. 1 is described by a single function, and the electromagnetic field in vacuum must be described either by six functions, three components of **E** and three components of **B**, or by four functions A^μ which, however, are not unique, or by the spinor functions ϕ_{ab}. In this section we shall discuss quite generally field equations for arbitrary fields from the standpoint of the action principle, which also guarantees automatically the relativistic invariance of the theory.

Let $\psi^\alpha(x)$ be a set of real or complex field quantities that may be scalar, vector, spinor, or tensor functions of the four-vector x. The Lagrangian density, which we shall denote from now on by L and refer to simply as Lagrangian,

$$L = L(\psi^\alpha(x), \psi^\alpha_{,\mu}(x), x^\mu), \tag{3.47}$$

is assumed to be a scalar and to depend, for the time being, on the first derivatives of $\psi^\alpha(x)$ only.

We consider the variational principle

$$\delta \int_R L\, dx = 0, \qquad dx \equiv d^4x = dx^0\, dx^1\, dx^2\, dx^3, \tag{3.48}$$

where R is a four-dimensional space-time region with a three-dimensional boundary B. We shall specify precisely the variation δ, and we shall consider several different variations.

(A) Variations with a Fixed Boundary *B*

In this case, clearly, the variational principle becomes

$$\delta \int_R L\, dx = \int_R \delta L\, dx = 0. \tag{3.49}$$

We can then write, from Eq. 3.47, the variation of L explicitly

$$\delta L = \frac{\partial L}{\partial \psi^\alpha}\,\delta\psi^\alpha + \frac{\partial L}{\partial \psi^\alpha_{,\mu}}\,\delta\psi^\alpha_{,\mu}.$$

The coordinates x^μ are not varied; δL is due entirely to the variation of the field quantities. We can rewrite δL in the form

$$\delta L = \left(\frac{\partial L}{\partial \psi^\alpha} - \partial_\mu \frac{\partial L}{\partial \psi^\alpha_{,\mu}}\right)\delta\psi^\alpha + \partial_\mu\left(\frac{\partial L}{\partial \psi^\alpha_{,\mu}}\,\delta\psi^\alpha\right). \tag{3.50}$$

We shall often use the following short notations

$$L_\alpha \equiv \frac{\partial L}{\partial \psi^\alpha}, \qquad L^\mu_\alpha \equiv \frac{\partial L}{\partial \psi^\alpha_{,\mu}}, \qquad \Lambda_\alpha \equiv L_\alpha - \partial_\mu L^\mu_\alpha. \tag{3.50$'$}$$

Then

$$\int_R \delta L\,dx = \int_R \Lambda_\alpha \delta\psi^\alpha\,dx + \int_R \partial_\mu(L^\mu_\alpha \delta\psi^\alpha)\,dx\,;$$

the last divergence term can be transformed by Gauss's theorem (see Chapter I, Sect. 8) into a 3-dimensional surface integral:

$$\int_R \delta L\,dx = \int_R \Lambda_\alpha \delta\psi^\alpha\,dx + \int_B d\sigma_\mu L^\mu_\alpha\,\delta\psi^\alpha. \tag{3.51}$$

Now, we further specify the variation δ in such a way that $\delta\psi^\alpha$ vanishes on the boundary B of R, but is completely arbitrary elsewhere. The last term in Eq. 3.51 then vanishes. Because $\delta\psi^\alpha$ are arbitrary in R, the action principle, Eq. 3.49, therefore implies the equations

$$\Lambda_\alpha = \frac{\partial L}{\partial \psi^\alpha} - \partial_\mu\left(\frac{\partial L}{\partial \psi^\alpha_{,\mu}}\right) = 0. \tag{3.52}$$

These equations will be taken as *field equations*. To summarize, we have shown the following result: Under a variation that does not change the region of integration, nor the coordinates, and such that the variation of the field quantities ψ^α vanish on the boundary of the integration region, the action principle, Eq. 3.49, gives the field equations, Eqs. 3.52. Physically, the actual field quantities make the action integral $\int L\,dx$ an extremum under the above variation. It is the philosophy of the Lagrangian field theory that this principle is valid for all physical systems. Therefore, the dynamics of the system is specified once the Lagrangian is given. Of course, the variational principle, does not tell us how to choose the Lagrangian. The choice of the proper Lagrangian is the more physical aspect of the theory. The mathematical part of the theory is to obtain and solve the field equations. Often one is not interested in the complete solutions of the partial differential equations. Again the physics will tell us what type of information is meaningful and how to obtain it. The situation here is quite analogous to other branches of physics. In mechanics the specific form of the forces, or the Hamiltonian, must be given in order to solve the equations of motion. In practice one knows the Lagrangian or the Hamiltonian of simple systems. One constructs complicated ones from their combinations or generalizations, as the case may be.

It is important to note here that the specification of a field by a Lagrangian is not unique. The Lagrangian determines the field equations up to an additive divergence term. For, if we add to L a term $\partial_\mu \Gamma^\mu(\psi, x)$, we get

$$\delta \int_R (L + \partial_\mu \Gamma^\mu)\, dx = \delta \int_R L\, dx + \delta \int_B d\sigma_\mu \Gamma^\mu = \delta \int L\, dx, \qquad (3.53)$$

because $\psi(x)$ is fixed on the boundary B under our variation. Notice that Γ^μ does not depend on the derivatives of $\psi(x)$. Therefore, we have the result that the set of Lagrangians $\{L + \partial_\mu \Gamma^\mu\}$, with arbitrary Γ^μ, refers to the same physical system.

If the field quantities $\psi^\alpha(x)$ are complex, we can take either their real and imaginary parts separately as independent field variables, or, alternatively, $\psi^\alpha(x)$ and $\psi^{*\alpha}(x)$ as independent quantities. In the latter case we will get similar field equations, Eqs. 3.52, for the complex conjugate $\psi^{*\alpha}(x)$.

Examples. Table I shows examples of Lagrangians and the field equations derived from them, according to Eq. 3.52. In all these examples the field quantities $\psi^\alpha(x)$ are ordinary real or complex numbers (c-numbers), although some of the fields mentioned (Dirac, Klein-Gordon fields) have quantum-mechanical importance.

Several forms of L are listed for the free electromagnetic field. These either differ from each other by a divergence term according to the above discussion, or use $F_{\mu\nu}$ or A_μ, or both, as the field quantities that are varied. In L_I, A_μ are taken to be the field variables and not $F_{\mu\nu}$, although L_I is written in terms of $F_{\mu\nu}$ to show its gauge invariant character. We have, then,

$$\frac{\partial L_I}{\partial A^\mu} = 0, \qquad \frac{\partial L}{\partial A^{\mu,\nu}} = F_{\mu\nu};$$

hence Eq. 3.52 gives $F^{\mu\nu}{}_{,\nu} = 0$, or $\Box^2 A_\mu = 0$, if we choose the *subsidiary conditions*

$$A^\mu_{,\mu} = 0$$

on the potentials.

Table I

Lagrangians and Equations of Motion for the Most Common Fields

Field	*Lagrangian*	*Field Equations*
Linear Chain	$\frac{1}{2}(\rho\dot{\psi}^2 - Y\psi^2{}_{,x})$	$\ddot{\psi} = \frac{Y}{\rho}\frac{\partial^2\psi}{\partial x^2}$
Real Scalar Field	$\frac{1}{2}(g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} - m^2\phi^2)$	$(\square^2 + m^2)\phi = 0$
Complex Scalar Field	$\frac{1}{2}(\phi^{,\mu}\phi^*{}_{,\mu} - m^2\phi\phi^*)$	$(\square^2 + m^2)\phi = 0; \quad (\square^2 + m^2)\phi^* = 0$
Free Electromagnetic Field	$L_{\text{I}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}(E^2 - B^2)$ $L_{\text{II}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(A^\mu{}_{,\mu})^2$ $L_{\text{III}} = -\frac{1}{2}A^\mu{}_{,\nu}A_\mu{}^{,\nu}$ $L_{\text{IV}} = \frac{1}{2}[A_\nu F^{\mu\nu}{}_{,\mu} - A_{\nu,\mu}F^{\mu\nu}] + \frac{1}{4}F^{\mu\nu}F_{\mu\nu}$	$F^{\mu\nu}{}_{,\nu} = 0$ $\square^2 A_\mu = 0$ $\square^2 A_\mu = 0$ $\square^2 A_\mu = 0$
Electromagnetic Field with an External Current	$-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}A_\mu j^\mu$	$F^{\mu\nu}{}_{,\nu} = -\frac{1}{c}j^\mu$
Dirac Field	$L_{\text{I}} = i\bar{\psi}\gamma^\mu\partial_\mu\psi + m\bar{\psi}\psi$ $L_{\text{II}} = \frac{1}{2}\bar{\psi}(i\gamma^\mu\overrightarrow{\partial_\mu} + m)\psi + \frac{1}{2}\bar{\psi}(-i\gamma^\mu\overleftarrow{\partial_\mu} + m)\psi$	$(i\gamma^\mu\partial_\mu + m)\psi(x) = 0; \quad i\bar{\psi}_{,\mu}\gamma^\mu - m\bar{\psi} = 0$ $(i\gamma^\mu\partial_\mu + m)\psi(x) = 0; \quad i\bar{\psi}_{,\mu}\gamma^\mu - m\bar{\psi} = 0$
Interacting Dirac and Maxwell Fields	$L_{\text{Dirac}} + L_{\text{Maxwell}} - \frac{e}{c}A_\mu(x)\bar{\psi}\gamma^\mu\psi$	$\square^2 A_\mu(x) = e\bar{\psi}\gamma_\mu\psi$ $\left[\gamma^\mu\left(i\partial_\mu - \frac{e}{c}A_\mu(x)\right) + m\right]\psi(x) = 0$

L_{II} is not gauge-invariant (Fermi-Lagrangian), but it gives the field equations $\Box^2 A_\mu = 0$ without subsidiary conditions. L_{III} differs from L_{II} by a divergence term $\frac{1}{2}\,\partial_\mu(A_\nu A^{\mu,\nu})$ if $A^{\mu}_{,\mu} = 0$. Finally in L_{IV} we consider both A_μ and $F^{\mu\nu}$ as independent field quantities (Schwinger-Lagrangian). The variation of L_{IV} with respect to both of these quantities gives

$$\Box^2 A_\mu = 0\,; \qquad F^{\mu\nu}_{,\nu} = 0,$$

and the connection between $F_{\mu\nu}$ and A_μ, namely

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}.$$

For the Dirac field a 4-component complex field quantity $\psi^\alpha(x)$, $\alpha = 1, 2, 3, 4$, and

$$\bar{\psi}^\alpha(x) = \psi^{\alpha *}(x)\gamma^0$$

are taken as the 8 independent field variables to be varied. Here the γ^μ are four 4 × 4 matrices satisfying

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}.$$

The two Lagrangians for the Dirac field differ by a divergence term $\frac{1}{2}\, i\partial_\mu(\bar{\psi}\gamma^\mu\psi)$.

The last example, the interaction of Dirac and Maxwell fields is discussed in the second part of the book.

(B) Variations Involving a Change of Boundary B

To distinguish from the previous case, we denote this variation by δ'. This variation involves a variation of the field quantities

$$\psi'(x) = \psi(x) + \delta\psi(x),$$

and a variation of the coordinates

$$x'^\mu = x^\mu + \delta x^\mu.$$

Let R' be the new varied region of integration. Then

$$\delta' \int_R L\, dx = \int_R \delta L\, dx + \int_{R'-R} L\, dx = \int_R \delta L\, dx + \int_B L\, d\sigma_\mu \delta x^\mu. \qquad (3.54)$$

This is illustrated in Fig. 4. The first term on the right-hand side has been evaluated already in Eq. 3.51. Thus,

$$\delta' \int_R L\, dx = \int_R \Lambda_\alpha \delta\psi^\alpha\, dx + \int_B d\sigma_\mu [L^\mu_\alpha \delta\psi^\alpha + L\delta x^\mu]. \qquad (3.55)$$

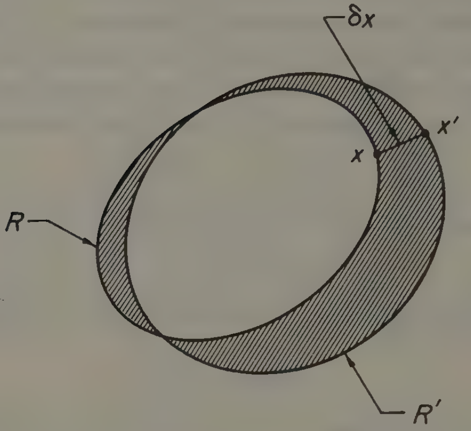


Fig. 4. Variation of the region of integration in the Action Principle $\int_R L\,dx$. The integral over the shaded region is equal to $\int_B d\sigma_\mu \delta x^\mu L$

Assuming now that the field equations are satisfied, i.e., $\Lambda_\alpha = 0$, we get

$$\delta' \int_R L\,dx = \int_B d\sigma_\mu [L^\mu_\alpha \delta\psi^\alpha + L\delta x^\mu]. \tag{3.56}$$

This equation relates the δ' variation of the action to a definite boundary integral. We now re-express $\delta\psi^\alpha$ in terms of variation of fields at the corresponding points under variation; i.e.,

$$\delta'\psi = \psi'(x') - \psi(x) = \psi'(x + \delta x) - \psi(x) = \delta\psi + \frac{\partial\psi}{\partial x^\mu}\,\partial x^\mu ;$$

$\delta\psi$ compares the two fields at the same space-time point, whereas $\delta'\psi$ compares them at the point x and the varied point x', respectively. In terms of $\delta'\psi$, Eq. 3.56 becomes

$$\delta' \int_R L\,dx = \int_B d\sigma_\mu [L^\mu_\alpha \delta'\psi^\alpha + (g^\mu_\nu L - \psi^\alpha_{,\nu} L^\mu_\alpha)\delta x^\nu] = 0. \tag{3.57}$$

Therefore, the new variation principle with the variation of the boundary gives the equation, after transforming Eq. 3.57 into a volume integral,

$$\partial_\mu [L^\mu_\alpha \delta'\psi^\alpha + (g^\mu_\nu L - \psi^\alpha_{,\nu} L^\mu_\alpha)\,\delta x^\nu] = 0. \tag{3.57'}$$

In the next section we use Eqs. 3.56 and 3.57 to discuss the conservation laws. Equation 3.57′ is used for an alternate derivation of the symmetric energy momentum tensor (see the solution of Exercise 10).

4. *Conservation Laws*

The principle of relativity has played a central rôle in our development so far. According to this principle, we are describing permanencies of nature independent of any coordinate frame. Closely related with this fact is the existence of certain quantities whose values do not change during the evolution of the system, the so-called *constants of motion*, These quantities are easy to observe or measure. Important physical concepts, such as energy, momentum, angular momentum, etc., have been introduced, because these quantities are constants of motion under special conditions, or for simple systems. A complicated physical system may be characterized by its constants of motion rather than by the detailed behavior of its parts. The importance of such *global* quantities is best known in thermodynamics.

In addition to the principle of relativity, we may have additional internal symmetry properties for physical systems. The laws or the equations of the system are then invariant under certain transformations corresponding to this symmetry. In turn these transformations give rise to the existence of conservation laws.

In this section we give a general discussion of the connection between the invariance and symmetry, on the one hand, and the conservation laws or constants of motion on the other, based on the action principle. We obtain first the integral expressions for the global quantities which are constants of motion, and then the equations of continuity which show what happens at each point, and finally their connection.

(A) Conservation Laws in Integral Form

In Eq. 3.56 we have calculated the variation of the action under a variation or a transformation of x^μ and $\psi^\alpha(x)$ simultaneously. Let us take the region R in this formula to be specified as follows. It is surrounded by two space-like surfaces σ_1 and σ_2 extending to infinity. (Fig. 5.) Also we make the important but plausible assumption that the fields $\psi^\alpha(x)$, their derivatives, L and $L_\alpha^\mu, \cdots$, all tend to zero at spatial infinity. Clearly, only then we can expect to find globally conserved quantities. We can then assume that B consists of σ_1 and σ_2 alone; the contribution of the surface parts at infinity can be neglected. Equation 3.56 becomes

$$\delta' \int_R L\,dx = \int_B d\sigma_\mu[L\delta x^\mu + L_\alpha^\mu \delta\psi^\alpha] \cong \int_{\sigma_1}^{\sigma_2} d\sigma_\mu[L\delta x^\mu + L_\alpha^\mu \delta\psi^\alpha]$$
$$= G(\sigma_2) - G(\sigma_1), \qquad \textbf{(3.58)}$$

where we have defined the *generator* of the variation δ' as

$$G(\sigma) = \int_\sigma d\sigma_\mu \left(L\delta x^\mu + \frac{\partial L}{\partial \psi^\alpha_{,\mu}} \delta\psi_\alpha \right). \tag{3.59}$$

Thus G is given by an integral over a three-dimensional surface; in particular, if we choose $d\sigma_\mu$ along the time-direction,

$$G(t) = c \int d^3x \left(L\delta t + \frac{\partial L}{\partial \psi^\alpha} \delta\psi^\alpha \right). \tag{3.59'}$$

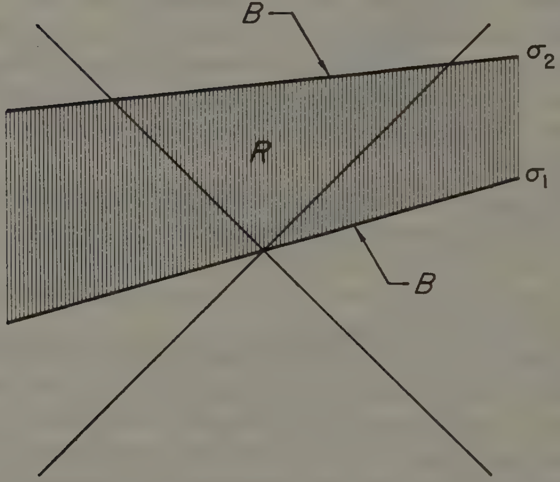


Fig. 5. Space-like surfaces σ_1 and σ_2, the four-dimensional volume R and the three-dimensional boundary B

Now let us assume that our system has an invariance or a symmetry property under the variation (or transformation) expressed by the fact that the variation of the action vanishes; i.e.,

$$\delta' \int_R L\, dx = 0. \tag{3.60}$$

Then it follows from Eq. 3.58 that $G(\sigma)$ does not depend on the space-like surface σ, nor on time; thus the generator $G(\sigma)$ of the variation is a constant of motion:

$$G = \text{constant}. \tag{3.61}$$

To summarize, if a variation $x' = x + \delta x$, $\psi'(x) = \psi(x) + \delta\psi(x)$, does not change the action $\int_R L\, dx$, then G given by Eq. 3.59 is a constant of the motion, and is the same for all coordinate frames.

If G is constant and invariant, it follows from Gauss's theorem[1] that the integrand of G is divergenceless:

$$\partial_\mu \left(L\delta x^\mu + \frac{\partial L}{\partial \psi^\alpha_{,\mu}} \delta\psi^\alpha \right) = 0. \tag{3.62}$$

This is a generalized equation of continuity. We have already seen a special example of it in electrodynamics, Eq. 3.31. There is always a one-to-one correspondence between a constant integral G and a local "equation of continuity". We shall discuss examples of this correspondence a little later.

It should be remarked that, although a divergence term $\partial_\mu \Gamma^\mu(\psi, x)$ added to the Lagrangian L does not change the field equations, it does change the value of the constant of motion to

$$G' = G + \int_B d\sigma_\mu\, \partial_\nu \Gamma^\nu(\psi, x)\, \delta x^\mu.$$

Finally, we should like to note here, why $\int L\, dx$ is called the *action* rather than Hamilton's principal function, as in mechanics. Returning to Eq. 3.57, we see that for a variation that leaves $\int L\, dx$ invariant *and* for which $\delta'\psi = \psi'(x') - \psi(x) = 0$, we get

$$\int_B d\sigma_\mu (\delta^\mu_\nu L - \psi^\alpha_{,\nu} L^\mu_\alpha)\, \delta x^\nu = 0,$$

or

$$\partial_\mu \overset{\circ}{T}{}^\mu{}_\nu \equiv \partial_\mu(-\delta^\mu_\nu L + \psi^\alpha_{,\nu} L^\mu_\alpha) = 0. \tag{3.63}$$

In analogy to classical mechanics $\overset{\circ}{T}{}^\mu{}_\nu$, defined by this equation, may be called the canonical *energy-momentum tensor*. This is exactly the procedure used in classical mechanics to obtain the energy conservation from the action,[2] which justifies denoting $\int L\, dx$ as the action quantity of our system. The energy-momentum tensor will be discussed in detail presently.

Now we discuss the important special cases of our variation δ' and the corresponding conservation laws.

1. Coordinate Transformations. Let δ' consist entirely of coordinate transformations, Lorentz transformations or more general coordinate transformations. A physical point P is labeled in two coordinate frames by x^μ and $x'^\mu = x^\mu + \delta x^\mu$, respectively. This transformation induces a change on $\psi^\alpha(x)$, although the field variables themselves are not changed.

---

[1] See Chapter I, Sect. 8.

[2] See, for example, H. Goldstein, *op. cit.*, p. 228.

At the same physical point P the field quantities in two coordinate frames are labeled by $\psi^{\alpha}(x)$ and $\psi'^{\alpha}(x')$, respectively. $\psi'^{\alpha}(x')$ is given according to the nature of the fields, i.e.; whether they are scalar, vector, spinor, tensor, etc. Let us write quite generally

$$\psi'^{\alpha}(x') = [\delta^{\alpha}_{\beta} + \Lambda^{\alpha}_{\beta}(L)]\psi^{\beta}(x), \tag{3.64}$$

$$\begin{aligned} &\Lambda^{\alpha}_{\beta} = 0, && \text{for a scalar field ;} \\ &(1+\Lambda)^{\alpha}_{\beta}, = L^{\alpha}{}_{\beta} && \text{for a vector field ;} \\ &(1+\Lambda)^{\alpha}_{\beta} = S(L)^{\alpha}_{\beta}, && \text{for a spinor field, } etc., \end{aligned}$$

and L^{α}_{β} is the Lorentz transformation (or possibly a more general coordinate transformation). On the other hand,

$$\psi'^{\alpha}(x') = \psi'^{\alpha}(x + \delta x) = \psi'^{\alpha}(x) + \frac{\partial \psi'^{\alpha}}{\partial x^{\mu}}\,\delta x^{\mu}. \tag{3.65}$$

Combining Eqs. 3.64 and 3.65, we get the induced variation on ψ:

$$\delta\psi^{\alpha} = \psi'^{\alpha}(x) - \psi^{\alpha}(x) = \Lambda^{\alpha}_{\beta}\psi^{\beta}(x) - \psi^{\alpha}_{,\mu}\,\delta x^{\mu}. \tag{3.65'}$$

Therefore, by Eq. 3.59, the generator of the coordinate transformations is given by

$$G(\sigma) = \int_{\sigma} d\sigma_{\mu}[L\delta x^{\mu} + L^{\mu}_{\alpha}(\Lambda^{\alpha}_{\beta}\psi^{\beta} - \psi^{\alpha}_{,\nu}\delta x^{\nu})]. \tag{3.66}$$

To further evaluate G, we now consider an infinitesimal Lorentz transformation

$$x'^{\mu} = x^{\mu} + \omega^{\mu}{}_{\nu}x^{\nu} + a^{\mu},$$

or

$$\delta x^{\mu} = \omega^{\mu}{}_{\nu}x^{\nu} + a^{\mu}, \tag{3.67}$$

where a^{μ} is a constant vector, and $\omega^{\mu\nu} = -\omega^{\nu\mu}$ is a constant antisymmetric tensor.[1] The quantities Λ^{α}_{β} in Eq. 3.64 depend only on $\omega_{\mu\nu}$. If we write

$$\Lambda^{\alpha}_{\beta} = \tfrac{1}{2}S^{\alpha\mu\nu}_{\beta}\,\omega_{\mu\nu}, \tag{3.68}$$

then

$$S^{\alpha\mu\nu}_{\beta} = -S^{\alpha\nu\mu}_{\beta}, \tag{3.69}$$

and

$$\begin{aligned} S^{\alpha\mu\nu}_{\beta} &= 0, && \text{if } \psi \text{ is a scalar field,} \\ &= g^{\alpha\mu}g^{\nu}_{\beta} - g^{\alpha\nu}g^{\mu}_{\beta}, && \text{if } \psi \text{ is a vector field,} \\ &= \tfrac{1}{4}(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu})^{\alpha}_{\beta}, && \text{if } \psi \text{ is } a \text{ 4-component spinor, } etc.^{2} \end{aligned}$$

---

[1] See Chapter I, Sect. 4.

[2] See Exercise 12.

We now rewrite the middle term in Eq. 3.66 in the form

$$L^{\mu}_{\alpha}\Lambda^{\alpha}_{\beta}\psi^{\beta} = \tfrac{1}{2}L^{\mu}_{\alpha}S^{\alpha\sigma\rho}_{\beta}\omega_{\sigma\rho}\psi^{\beta} \equiv f^{\mu\sigma\rho}\omega_{\sigma\rho}\,. \tag{3.70}$$

Now, from Eq. 3.67,

$$\frac{\partial}{\partial x^{\rho}}(\delta x_{\sigma}) = \omega_{\sigma\rho} = -\frac{\partial}{\partial x^{\sigma}}(\delta x_{\rho})\,;$$

hence

$$f^{\mu\sigma\rho}\omega_{\sigma\rho} = -\frac{\partial}{\partial x^{\sigma}}(f^{\mu\sigma\rho}\delta x_{\rho}) + \frac{\partial}{\partial x^{\sigma}}(f^{\mu\sigma\rho})\delta x_{\rho}\,.$$

The first term, when integrated over the space-like surface σ, gives

$$\begin{aligned}\int d\sigma_{\mu}\frac{\partial}{\partial x^{\sigma}}(f^{\mu\sigma\rho}\delta x_{\rho}) &= \tfrac{1}{2}\textstyle\int d\sigma_{\mu}[(f^{\mu\sigma\rho}\delta x_{\rho})_{,\sigma} + (f^{\mu\sigma\rho}\delta x_{\rho})_{,\sigma}]\\ &= \tfrac{1}{2}\textstyle\int d\sigma_{\mu}(f^{\mu\sigma\rho}\delta x_{\rho})_{,\sigma} - \tfrac{1}{2}\int d\sigma_{\sigma}(f^{\mu\sigma\rho}\delta x_{\rho})_{,\mu}\,,\end{aligned} \tag{3.71}$$

where in the second term we have interchanged the dummy indices μ and σ and then used the antisymmetry of $f^{\mu\sigma\rho}$ in the indices μ and σ, which will be justified below. The contribution of this term to G is a constant independent of whether or not we have Lorentz invariance. For if we evaluate Eq. 3.71 between two space-like surfaces and use Gauss's theorem,

$$\begin{aligned}\left(\int_{\sigma_2} - \int_{\sigma_1}\right)[d\sigma_{\mu}(f^{\mu\sigma\rho}\delta x_{\rho})_{,\sigma} - d\sigma_{\sigma}(f^{\mu\sigma\rho}\delta x_{\rho})_{,\mu}] = \int_{V} dx[&(f^{\mu\sigma\rho}\delta x_{\rho})_{,\sigma\mu}\\ &- (f^{\mu\sigma\rho}\delta x_{\rho})_{,\mu\sigma}] = 0.\end{aligned}$$

Therefore we can omit this constant number and obtain

$$G(\sigma) = \int_{\sigma} d\sigma_{\mu}[L\delta x^{\mu} + f^{\mu\sigma\rho}{}_{,\sigma}\delta x_{\rho} - L^{\mu}_{\alpha}\psi^{\alpha}_{,\nu}\delta x^{\nu}] = -\int_{\sigma} d\sigma_{\mu}T^{\mu\nu}\delta x_{\nu}\,, \tag{3.72}$$

where we have defined the so-called *energy-momentum tensor* $T^{\mu\nu}$ of the fields,

$$T^{\mu\nu} = L^{\mu}_{\alpha}\psi^{\alpha,\nu} - f^{\mu\sigma\nu}{}_{,\sigma} - Lg^{\mu\nu} \tag{3.73}$$

Before we discuss the meaning of $T^{\mu\nu}$, we must prove that $f^{\mu\sigma\nu}$ is antisymmetric in μ and σ. From its defining equation, Eq. 3.70, we obtain

$$f^{\mu\sigma\rho} = \tfrac{1}{2}(L^{\mu}_{\alpha}S^{\alpha\sigma\rho}_{\beta} + L^{\sigma}_{\alpha}S^{\alpha\rho\mu}_{\beta} - L^{\rho}_{\alpha}S^{\alpha\mu\sigma}_{\beta})\psi^{\beta}\,; \tag{3.74}$$

indeed, if we multiply both sides of this equation with $\omega_{\sigma\rho}$ and sum, we get

$$f^{\mu\sigma\rho}\omega_{\sigma\rho} = \tfrac{1}{2}(L^{\mu}_{\alpha}S^{\alpha\sigma\rho}_{\beta}\omega_{\sigma\rho} + L^{\sigma}_{\alpha}S^{\alpha\rho\mu}_{\beta}\omega_{\sigma\rho} - L^{\rho}_{\alpha}S^{\alpha\mu\sigma}_{\beta}\omega_{\sigma\rho})\psi^{\beta}.$$

If we change the dummy indices ρ and σ in the last term and make use of the fact that both $\omega_{\sigma\rho}$ and $S^{\mu\sigma}$ are antisymmetric, we see that the last two terms cancel and we get back the correct defining equation, Eq. 3.70. From Eq. 3.74 and the antisymmetry of $S^{\mu\sigma}$ we finally obtain the desired result:

$$f^{\mu\sigma\rho} = -f^{\sigma\mu\rho}. \tag{3.75}$$

We shall discuss now the physical significance of the energy-momentum tensor $T^{\mu\nu}$. Let us use explicitly the fact that we are considering an infinitesimal Lorentz transformation, and that δx^μ is given by Eq. 3.67. Hence the generator $G(\sigma)$ becomes

$$G(\sigma) = -\int_\sigma d\sigma_\mu T^{\mu\nu}(\omega_{\nu\sigma}x^\sigma + a_\nu).$$

Or, if we introduce the quantities

$$P^\nu \equiv \int_\sigma d\sigma_\mu T^{\mu\nu}; \tag{3.76}$$

$$J^{\nu\sigma} \equiv \int_\sigma d\sigma_\mu M^{\mu\nu\sigma} \equiv \int_\sigma d\sigma_\mu(T^{\mu\nu}x^\sigma - T^{\mu\sigma}x^\nu), \tag{3.77}$$

$G(\sigma)$ can be written in the form

$$G(\sigma) = -a_\nu P^\nu - \tfrac{1}{2}\omega_{\nu\sigma}J^{\nu\sigma}. \tag{3.78}$$

The quantity P^ν is the generator corresponding to translations and $J^{\mu\nu}$ that corresponding to "rotations". If for the system of fields we are considering, the Lagrangian is invariant under the inhomogeneous Lorentz transformations, these generators will be constants of motion. They must, therefore, be interpreted as the *total* energy and angular momentum of the system, respectively.

Because L is invariant, the derivatives $L^\mu_\alpha = \partial L/\partial\psi^\alpha{}_{,\mu}$ have definite transformation properties. We see from its defining equation, Eq. 3.73, that $T^{\mu\nu}$ is indeed a tensor. Therefore, P^μ is a four-vector and $J^{\mu\nu}$ is a tensor.

If $G(\sigma)$ is a constant of motion, we obtain from Gauss's theorem and the vanishing of $T^{\mu\nu}$ at spatial infinity the equations of continuity

$$T^{\mu\nu}{}_{,\mu} = 0, \tag{3.79}$$

$$M^{\mu\nu\sigma}{}_{,\mu} = 0. \tag{3.80}$$

These equations in turn imply that $T^{\mu\nu}$ is symmetric, for

$$M^{\mu\nu\sigma}{}_{;\mu} = T^{\mu\nu}{}_{,\mu} x^{\sigma} + T^{\mu\nu}\delta^{\sigma}_{\mu} - T^{\mu\sigma}{}_{,\mu} x^{\nu} - T^{\mu\sigma}\delta^{\nu}_{\mu},$$

hence[1]

$$T^{\sigma\nu} = T^{\nu\sigma}. \tag{3.81}$$

We see that the variation introduced at the beginning of this section automatically yields a *symmetric* energy-momentum tensor. Notice that the previously introduced, so-called "canonical energy-momentum" tensor, Eq. 3.63, is not symmetric. We have the relation

$$T^{\mu\nu} = \overset{\circ}{T}{}^{\mu\nu} - f^{\mu\sigma\nu}{}_{,\sigma}. \tag{3.82}$$

The two tensors become identical for scalar fields; their difference is due to the fact that in calculating $T^{\mu\nu}$ we have taken into account the transformation properties of the underlying fields.

The infinitesimal Lorentz group has ten parameters. The corresponding ten generators P^{μ} and $J^{\mu\nu}$ are identical with those of the Lorentz group discussed in Chapter I, Sect. 5. To evaluate P^{μ} explicitly, one can conveniently choose a frame such that the surface $t =$ constant is the space-like surface. Then

$$P^{\mu} = \int_V T^{0\mu}\, d^3x;$$

in particular, the total energy is the space-integral of T^{00}. In the particular coordinate system considered, the equation of continuity $T^{\mu\nu}{}_{,\nu} = 0$ implies

$$\begin{aligned} \frac{\partial}{\partial x^0} \int_V T^{\mu 0}\, d^3x &= -\int_V T^{\mu k}{}_{,k}\, d^3x; \qquad k = 1, 2, 3, \\ &= -\int_S T^{\mu k}\, df_k, \end{aligned}$$

where V is a three-dimensional volume and S is its surface. The last integral vanishes if $T^{\mu k}$ vanishes at spatial infinity. Hence $\int_V T^{\mu 0} d^3x = P^{\mu}$ is both a constant and a Lorentz invariant.

2. Change of Field Variables. Here we consider an intrinsic variation of the field quantities so that no coordinate transformation is involved:

$$\delta x^{\mu} = 0; \qquad \bar{\psi}(x) = \psi(x) + \delta_0 \psi.$$

---

[1] We see then from Eq. 3.73 that

$$L_{\alpha}{}^{\mu}\psi^{\alpha,\nu} - L_{\alpha}{}^{\nu}\psi^{\alpha,\mu} = f^{\mu\sigma\nu}{}_{,\sigma} - f^{\nu\sigma\mu}{}_{,\sigma}. \tag{3.81$'$}$$

From Eq. 3.59 the generator corresponding to this variation is given by

$$G(\sigma) = \int_\sigma d\sigma_\mu L^\mu_\alpha \delta_0 \psi^\alpha = \int_\sigma d\sigma_\mu j^\mu, \tag{3.83}$$

where we have introduced a certain four-vector j^μ corresponding to the variation considered:

$$j^\mu = L^\mu_\alpha \delta_0 \psi^\alpha. \tag{3.84}$$

Now if this particular variation of ψ leaves $\int_R L dx$ invariant, then $G(\sigma)$ given by Eq. 3.83 is again a constant of motion, and assuming that the fields fall off to zero at spatial infinity, we get the equation of continuity

$$j^\mu_{,\mu} = 0. \tag{3.85}$$

We shall see that the gauge transformations and the corresponding continuity equation for the current belong to this category. The generator in Eq. 3.83 may be interpreted as the total generalized *charge* of the system.

3. Coordinate Transformations and Intrinsic Variations. In the general case of a simultaneous coordinate transformation plus an intrinsic change of the field variables, $\delta\psi$ in the fundamental equation, Eq. 3.59, consists of two parts, one comes from coordinate transformations (induced $\delta\psi$-variation) and is given by Eq. 3.65′, and the other is the intrinsic variation $\delta_0\psi$. Hence

$$\delta\psi = \Lambda^\alpha_\beta \psi^\beta - \psi^\alpha_{,\mu} \delta x^\mu + \delta_0 \psi,$$

and the generator is given by

$$\begin{aligned} G(\sigma) &= \int_\sigma d\sigma_\mu [L \delta x^\mu + L^\mu_\alpha (\Lambda^\alpha_\beta \psi^\beta - \psi^\alpha_{,\nu} \delta x^\nu + \delta_0 \psi)] \\ &= -\int_\sigma d\sigma_\mu [T^{\mu\nu} \delta x_\nu - L^\mu_\alpha \delta_0 \psi^\alpha], \end{aligned} \tag{3.86}$$

which is the sum of Eq. 3.83 and the generalized charge.

(B) Conservation Laws in Differential Form

In this section we shall derive the conservation laws in differential form directly from the Lagrangian without using the action integral. We go back to Eq. 3.50 and assume that the field equations are satisfied. Then the variation of L is given simply by

$$\delta L = \partial_\mu (L^\mu_\alpha \delta \psi^\alpha). \tag{3.87}$$

This equation connects δL with an arbitrary variation $\delta\psi$ *consistent* with the equations of motion; i.e., variations which leave the equation of motion invariant.

If, in addition to the equation of motion, the Lagrangian itself is invariant; i.e., $\delta L = 0$, we get immediately the equation of continuity

$$\partial_\mu \left(L^\mu_\alpha \delta\psi^\alpha\right) = 0. \tag{3.88}$$

For an intrinsic variation $\delta_0\psi$, this is precisely Eq. 3.85 with j^μ given by Eq. 3.84.

Now let us consider the infinitesimal inhomogeneous Lorentz transformations, Eq. 3.67, and evaluate both sides of Eq. 3.87 under these transformations. We have

$$\begin{aligned}\delta L = L'(x) - L(x) &= \frac{\partial L}{\partial x^\mu}\,\delta x^\mu = -\frac{\partial L}{\partial x^\mu}(\omega^{\mu\nu}x_\nu + a^\mu)\\ &= -\frac{\partial L}{\partial x^\mu}\,a^\mu - \tfrac{1}{2}\omega^{\mu\nu}\left(x_\nu\frac{\partial L}{\partial x^\mu} - x_\mu\frac{\partial L}{\partial x^\nu}\right).\end{aligned}$$

The variation of $\delta\psi$ under Eq. 3.67 has been evaluated already in Eq. 3.65′:

$$\begin{aligned}\delta\psi^\alpha = \psi'^\alpha(x) - \psi^\alpha(x) &= \tfrac{1}{2}S^{\alpha\mu\nu}_\beta\,\omega_{\mu\nu}\psi^\beta - \psi^\alpha_{,\mu}\delta x^\mu\\ &= -\psi^\alpha_{,\mu}a^\mu + \tfrac{1}{2}\omega_{\mu\nu}[S^{\alpha\mu\nu}_\beta\psi_\beta - (x^\mu\psi^{\alpha,\nu} - x^\nu\psi^{\alpha,\mu})].\end{aligned}$$

If we insert these expressions into Eq. 3.87, we obtain

$$\begin{aligned}-\frac{\partial L}{\partial x^\mu}\,a^\mu - \tfrac{1}{2}\omega^{\mu\nu}\left(x_\nu\frac{\partial L}{\partial x^\mu} - x_\mu\frac{\partial L}{\partial x^\nu}\right) + \partial_\mu\{L^\mu_\alpha\psi^\alpha_{,\nu}a^\nu\\ - \tfrac{1}{2}L^\mu_\alpha\omega_{\nu\sigma}[S^{\alpha\nu\sigma}_\beta\psi_\beta - (x^\nu\psi^{\nu,\sigma} - x^\sigma\psi^{\alpha,\nu})]\} = 0,\end{aligned}$$

or, using Eq. 3.70,

$$\begin{aligned}a^\nu\partial_\mu(-L\delta^\mu_\nu + L^\mu_\alpha\psi^\alpha_{,\nu}) + \tfrac{1}{2}\omega^{\sigma\rho}\partial_\mu[-\delta^\mu_\sigma x_\rho L + \delta^\mu_\rho x_\sigma L - f^{\mu\sigma\rho} + f^{\mu\rho\sigma}\\ + L^\mu_\alpha(x^\rho\psi^{\alpha,\sigma} - x^\sigma\psi^{\alpha,\rho})] = 0,\end{aligned}$$

or

$$a_\nu\mathring{T}^{\mu\nu}{}_{,\mu} + \tfrac{1}{2}\omega^{\sigma\rho}\mathring{M}^{\mu\sigma\rho}{}_{,\mu} = 0, \tag{3.89}$$

where $\mathring{T}^{\mu\nu}$ is the canonical, unsymmetric energy-momentum tensor

$$\mathring{T}^{\mu\nu} = L^\mu_\alpha\psi^{\alpha,\nu} - Lg^{\mu\nu}, \tag{3.90}$$

and we have introduced

$$\mathring{M}^{\mu\sigma\rho} = f^{\mu\sigma\rho} - f^{\mu\rho\sigma} + (\mathring{T}^{\mu\sigma}x^\rho - \mathring{T}^{\mu\rho}x^\sigma). \tag{3.91}$$

Because a_ν and $\omega^{\sigma\rho}$ are arbitrary, the equation of continuity follows from Eq. 3.89:

$$\mathring{T}^{\mu\nu}{}_{,\mu} = 0, \qquad \text{and} \qquad \mathring{M}^{\mu\sigma\rho}{}_{,\mu} = 0. \tag{3.92}$$

Therefore, this method leads first to the unsymmetrical $\mathring{T}^{\mu\nu}$ rather than to $T^{\mu\nu}$, as in the previous case. By Eq. 3.82 the difference between $T^{\mu\nu}$ and $\mathring{T}^{\mu\nu}$ is a divergence, $T^{\mu\nu} = \mathring{T}^{\mu\nu} - f^{\mu\sigma\nu}{}_{,\sigma}$; hence Eq. 3.92 implies that

$$T^{\mu\nu}{}_{,\mu} = 0,$$

and *vice versa*, because $f^{\mu\sigma\nu}$ is antisymmetric in the first two indices. Furthermore, the energy-momentum four-vectors evaluated from $T^{\mu\nu}$ or $\mathring{T}^{\mu\nu}$ differ by a fixed constant, because

$$\int_{\sigma_1}^{\sigma_2} d\sigma_\mu f^{\mu\sigma\nu}{}_{,\sigma} = \int_R dx f^{\mu\sigma\nu}{}_{,\sigma\mu} = 0.$$

We also show that, if Eq. 3.92 holds, $T^{\mu\nu}$ so defined must be symmetric. From $\mathring{M}^{\mu\sigma\rho}{}_{,\mu} = 0$ we get

$$\mathring{T}^{\mu\nu} - \mathring{T}^{\nu\mu} = f^{\sigma\nu\mu}{}_{,\sigma} - f^{\sigma\mu\nu}{}_{,\sigma};$$

on the other hand,

$$T^{\mu\nu} - T^{\nu\mu} = \mathring{T}^{\mu\nu} - \mathring{T}^{\nu\mu} - f^{\sigma\nu\mu}{}_{,\sigma} - f^{\sigma\mu\nu}{}_{,\sigma}.$$

Combining the last two equations, we see that $T^{\mu\nu} = T^{\nu\mu}$. Q. E. D.

Conversely, these results may be used to symmetrize the canonical energy-momentum tensor and to find the explicit form of $f^{\mu\nu\sigma}$ that should coincide with our previous derivation.[1] Thus, starting from $\mathring{T}^{\mu\nu}$ we can re-establish the symmetry of the energy-momentum tensor. Let

$$T^{\mu\nu} - \mathring{T}^{\mu\nu} = C^{\sigma\mu\nu}{}_{,\sigma}.$$

We want to show that $C^{\sigma\mu\nu} = f^{\sigma\mu\nu}$, if T has to be symmetric and also satisfy an equation of continuity. The latter requirement implies that

$$C^{\sigma\mu\nu} = -C^{\mu\sigma\nu},$$

and the former implies that

$$\mathring{T}^{\mu\nu} - \mathring{T}^{\nu\mu} = (C^{\sigma\nu\mu} - C^{\sigma\mu\nu})_{,\sigma}.$$

The left-hand side of this equation is known from Eq. 3.90. Let

$$\mathring{T}^{\mu\nu} - \mathring{T}^{\nu\mu} = L^{\mu}_{\alpha}\psi^{\alpha,\nu} - L^{\nu}_{\alpha}\psi^{\alpha,\mu} \equiv B^{\sigma\nu\mu}{}_{,\sigma}.$$

---

[1] F. Belinfante, *Physica*, 6 (1939), 887.

Hence it is sufficient to require that

$$B^{\sigma\nu\mu} = C^{\sigma\nu\mu} - C^{\sigma\mu\nu}.$$

A solution of this equation satisfying $C^{\sigma\mu\nu} = -C^{\mu\sigma\nu}$ is

$$C^{\sigma\nu\mu} = \tfrac{1}{2}(B^{\sigma\mu\nu} - B^{\nu\mu\sigma} + B^{\sigma\nu\mu}) = -C^{\nu\sigma\mu}.$$

Then, from Eqs. 3.74 and 3.81′, the desired results follow.

We see that we can formally symmetrize the canonical energy-momentum tensor by adding to it a suitable divergence term. However, the direct derivation of a symmetric energy-momentum tensor $T^{\mu\nu}$ from general principles is the more correct and fundamental approach. Another derivation of the symmetric $T^{\mu\nu}$ from the variational principle involving a change of the boundary is treated in Sect. 6, Exercise 10.

Returning to our starting point, Eq. 3.50, we see that for variations that leave the form of the Lagrangian invariant, $\delta L = 0$, but not the field equations, one obtains a generalization of Eq. 3.88,

$$\Lambda_\alpha \delta\psi^\alpha + \partial_\mu(L^\mu_\alpha \delta\psi^\alpha) = 0.$$

(C) Lagrangians Not Invariant Under Translations

The Lagrangian density L is always assumed to be a scalar with respect to homogeneous Lorentz transformations. But L may not be invariant under translations. An example, which we shall consider later in detail, is the case of an electromagnetic field interacting with a *given* external current.

In this case L has an explicit dependence on x^μ, and we get an additional term in Eq. 3.50. It becomes

$$\delta L = \Lambda_\alpha \delta\psi^\alpha + \partial_\mu(L^\mu_\alpha \delta\psi^\alpha) + \left.\frac{\partial L}{\partial x^\mu}\right|_{\text{expl}} \delta x^\mu\,; \tag{3.93}$$

the last term is due to the explicit dependence of L on x^μ. Consequently, for inhomogeneous Lorentz transformations, we obtain

$$\overset{\circ}{T}{}^{\mu\nu}{}_{,\mu} = -\left.\frac{\partial L}{\partial x^\mu}\right|_{\text{expl}}, \tag{3.94}$$

but

$$\overset{\circ}{M}{}^{\mu\sigma\rho}{}_{,\mu} = 0.$$

Clearly the energy and the momentum of the system is not conserved, but the total angular momentum still is.

(D) Explicit Form of Conserved Quantities

In Table II we have evaluated the energy-momentum tensor $T^{\mu\nu}$ and energy-momentum four-vector P^{μ} for a number of simple fields considered in this book, using the general theory of the present section. (See also Table I for the corresponding Lagrangians and the equations of motion.) The evaluation of $J^{\mu\nu}$ from $T^{\mu\nu}$ is left to the reader. (See Exercise 13.)

Some of these fields also admit an intrinsic variation of ψ, such as gauge transformations. In this case we also give the corresponding generator G and the (current) density j^{μ}.

For the scalar field the spin term $f^{\mu\sigma\rho}$ is zero, and we find

$$T^{\mu\nu} = \mathring{T}^{\mu\nu} = \phi^{,\mu}\phi^{,\nu} - g^{\mu\nu}L = \phi^{,\mu}\phi^{,\nu} - \tfrac{1}{2}g^{\mu\nu}(\phi^{,\sigma}\phi_{,\sigma} - m^2\phi^2).$$

The corresponding value of the energy-momentum vector P^{μ} in a particular frame is given in Table II ($P^{\mu} = \int d^3x T^{0\mu}$).

For the complex scalar field we find from the Lagrangian given in Table I that

$$T^{\mu\nu} = T^{\mu\nu}(\phi_1) + T^{\mu\nu}(\phi_2),$$

where ϕ_1 and ϕ_2 are the hermitian fields defined by

$$\phi_1 = \frac{1}{\sqrt{2}}(\phi + \phi^*), \qquad \phi_2 = \frac{1}{\sqrt{2}}(\phi - \phi^*).$$

Equation 3.95 can also be written in terms of ϕ and ϕ^*:

$$T^{\mu\nu} = \phi^{*,\mu}\phi^{,\nu} + \phi^{,\mu}\phi^{*,\nu} - g^{\mu\nu}L.$$

The complex scalar field admits also a gauge transformation. The Lagrangian $L(\phi_1) + L(\phi_2)$ is invariant under the transformation

$$\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix} \to (I + i\varepsilon)T\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix},$$

where T is a 2×2 matrix. Because ϕ_1 and ϕ_2 are real, T must be antisymmetric and imaginary:

$$T = \begin{pmatrix}0 & -i \\ i & 0\end{pmatrix}.$$

This then corresponds to the following transformation of the complex field:

$$\phi \to \phi + i\varepsilon\phi\,;$$

$$\phi^* \to \phi^{\approx} - i\varepsilon\phi^*,$$

as can also be verified directly by the invariance of the Lagrangian written in terms of ϕ and ϕ^*.

Table II
Conserved Quantities for Scalar and Electromagnetic Fields

	Scalar Field ϕ	*Complex Scalar Field* $\phi_1 = \frac{1}{\sqrt{2}}(\phi + \phi^*)$ $\phi_2 = \frac{-i}{\sqrt{2}}(\phi - \phi^*)$	*Free Electromagnetic Field*
$T^{\mu\nu}$	$\phi^{,\mu}\phi^{,\nu} - \frac{1}{2}g^{\mu\nu}(\phi^{,\sigma}\phi_{,\sigma} - m^2\phi^2)$	$T^{\mu\nu}(\phi_1) + T^{\mu\nu}(\phi_2)$	$F^{\sigma\mu}F_{\sigma}{}^{\nu} + \frac{1}{4}g^{\mu\nu}F^{\sigma\rho}F_{\sigma\rho}$ $T^{oo} = \frac{1}{2}(E^2 + B^2);\ T^{ok} = (\mathbf{E} \times \mathbf{B})_k$
P^{μ}	$\int d^3x\left[\frac{1}{c^2}\dot{\phi}\phi^{,\mu} - \frac{1}{2}\left(\frac{\dot{\phi}^2}{c^2} - (\nabla\phi)^2 - m^2\phi_2\right)\right]$	$P^{\mu}(\phi_1) + P^{\mu}(\phi_2)$	$P^o = \int d^3xT^{oo};\ P^k = \int' d^3xT^{ok}$
j^{μ}	—	$e(\phi^{,\mu}\phi^* - \phi\phi^{*,\mu})$	$F^{\alpha\mu}\Lambda_{,\alpha}$
G	—	$e\int d^3x(\dot{\phi}\phi^* - \phi\dot{\phi}^*)$	$\int d^3xF^{\alpha o}\Lambda_{,\alpha}$

The generator of the transformation, Eq. 3.97′, is given by

$$G = \int d\sigma_\mu L^\mu_\alpha \delta\psi^\alpha = \int d\sigma_\mu j^\mu,$$

and we find

$$j^\mu = e(\phi^{,\mu}\phi^* - \phi\phi^{*,\mu}).$$

The invariance of L implies that

$$j^\mu_{,\mu} = 0.$$

Physically j^μ is the current density and G the total charge of the Klein-Gordon particles. Therefore the one-component real scalar field describes only neutral particles.

The free electromagnetic field with the Lagrangian

$$L = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}$$

gives $\partial L/\partial A^{\mu,\nu} = F^{\mu\nu}$ and, according to Eq. 3.90,

$$\overset{\circ}{T}{}^{\mu\nu} = -F^{\sigma\mu}A_\sigma{}^{,\nu} + \tfrac{1}{4}g^{\mu\nu}F^{\sigma\rho}F_{\sigma\rho},$$

which is neither symmetric nor gauge invariant. We now symmetrize this canonical energy-momentum tensor according to the general principles discussed in the previous section. We obtain

$$f^{\mu\sigma\nu} = F^{\sigma\mu}A^\nu,$$

and

$$f^{\mu\sigma\nu}_{,\sigma} = A^\nu{}_{,\sigma}F^{\sigma\mu};$$

consequently,

$$T^{\mu\nu} = F^{\sigma\mu}A_\sigma{}^{,\nu} - A^\nu{}_{,\sigma}F^{\sigma\mu} - g^{\mu\nu}L = F^{\sigma\mu}F^\nu{}_\sigma - g^{\mu\nu}L,$$

which is obviously gauge invariant and symmetric.

It is instructive to write out $T^{\mu\nu}$ in terms of the *energy density* H, the *Poynting vector*

$$\mathbf{S} = c(\mathbf{E} \times \mathbf{B}),$$

and momentum density

$$\mathbf{G} = \frac{1}{c^2}\mathbf{S},$$

in the form

$$T^{\mu\nu} = \begin{pmatrix} H & \frac{1}{c}S_1 & \frac{1}{c}S_2 & \frac{1}{c}S_3 \\ cG_1 & & & \\ cG_2 & & T^{ik} & \\ cG_3 & & & \end{pmatrix}. \tag{3.95}$$

The equations $T^{\mu\nu}{}_{,\nu} = 0$ take then the more familiar forms

$$\frac{\partial H}{\partial t} + \nabla \cdot \mathbf{S} = 0\,;$$

$$\frac{1}{c}\frac{\partial}{\partial t} G_j + \frac{\partial}{\partial x^k} T^{jk} = 0\,; \qquad j = 1, 2, 3,$$

and express the continuity equation for energy and momentum.

The free electromagnetic field also admits a gauge transformation for which $\bar{\delta}\psi^{\alpha} = \Lambda^{,\alpha}$. The corresponding constant of motion is given, according to our general theory, by

$$G = \int d\sigma_{\mu} F^{\sigma\,\mu} \Lambda_{,\sigma}\,,$$

which corresponds to the continuity equation:

$$\partial_{\mu}(F^{\sigma\,\mu}\Lambda_{,\mu}) = 0.$$

5. *Canonical Form of Field Equations*

In this section we discuss the Hamiltonian form of the equation of motion for fields, generalizing the corresponding discussion for particles. (Chapter II, Sect. 3.) The distinction is that instead of the one independent variable s, the proper time, we have four independent variables x^{μ}. The Lagrangian density is a function of the field variables ψ^{α} and their derivatives: $L = L(\psi^{\alpha}, \psi^{\alpha}_{,\mu}, x)$. We define the conjugate field variables (or momenta) by

$$\Pi^{\mu}_{\alpha} \equiv \partial L/\partial \psi^{\alpha}_{,\mu} = L^{\mu}_{\alpha}\,, \tag{3.96}$$

and introduce a scalar function

$$C = \Pi^{\mu}_{\alpha}\psi^{\alpha}_{,\mu} - L, \tag{3.97}$$

or a tensor which is the canonical energy-momentum tensor already introduced:

$$\mathring{T}^{\mu\nu} = \Pi^{\mu}_{\alpha}\psi^{\alpha,\nu} - g^{\mu\nu}L. \tag{3.98}$$

Provided that we can solve in Eq. 3.96 $\psi^{\alpha}{}_{,\mu}$ as a function of Π^{μ}_{α}, ψ and x we can express C or $\mathring{T}^{\mu\nu}$ as a function of ψ^{α} and Π^{μ}_{α}:

$$C = C(\psi_{\alpha}\,, \Pi^{\mu}_{\alpha}\,, x).$$

If we now evaluate the derivatives of C with respect to ψ_α (or Π^μ_α) by fixed x and Π^μ_α (or ψ_α), we find that

$$\frac{\partial C}{\partial \Pi^\mu_\alpha} = \psi^\alpha_{,\mu} ;$$

$$\frac{\partial C}{\partial \psi^\alpha} = -\frac{\partial L}{\partial \psi^\alpha} = -\Pi^\mu_{\alpha,\mu} , \tag{3.99}$$

where in the last step we have made use of the field equations, Eqs. 3.52. Similarly, with $\mathring{T}^{\mu\nu}$:

$$\frac{\partial \mathring{T}^{\mu\nu}}{\partial \Pi^\mu_\alpha} = \psi^{\alpha,\nu} ;$$

$$\frac{\partial \mathring{T}^{\mu\nu}}{\partial \psi^\alpha} = -g^{\mu\nu}\frac{\partial L}{\partial \psi^\alpha} = -g^{\mu\nu}\Pi^\sigma_{\alpha,\sigma} . \tag{3.100}$$

Equations 3.99 and 3.100 are the *formal covariant Hamilton equations* of field theory. The success of the method lies—we repeat—in the possibility that one can solve $\psi^\alpha_{,\mu}$ in terms of Π^μ_α in Eq. 3.96. Table III shows some examples where this method works. The Lagrangians are given in Table I. In the case of the electromagnetic field, we must use the Lagrangian L_{III}. For the others one cannot solve $A^\mu_{,\nu}$ in terms of Π^μ_ν and A_μ. The table gives C and $\mathring{T}^{\mu\nu}$. The first set of the canonical equations are identically satisfied. The second set gives the equations of motion as shown in Table I.

The covariant quantities C and $\mathring{T}^{\mu\nu}$ used as "Hamiltonians" above do not have the meaning of energy. Similarly, in Chapter 2, Sect. 3, we have first introduced covariant "Hamiltonians" and then showed that certain energy expressions can be used as *bona fide* Hamiltonians. In field theory, also, the standard and useful practice is to use *noncovariant canonical equations* with the energy density being equal to *Hamiltonian density* $\mathscr{H}$,

$$\mathscr{H} \equiv \mathring{T}^{00} . \tag{3.101}$$

The canonical equations are now given, remembering that

$$\mathring{T}^{00} = \Pi_{\alpha 0}\psi^{\alpha,0} - L,$$

by

$$\frac{\partial \mathscr{H}}{\partial \Pi_{\alpha 0}} = \psi^{\alpha,0} \equiv \dot{\psi}^\alpha , \tag{3.102}$$

and

$$\frac{\partial \mathscr{H}}{\partial \psi^\alpha} = -\frac{\partial L}{\partial \psi^\alpha} . \tag{3.102'}$$

Table III

Conjugate Variables and Covariant "Hamiltonians"

Field	*Conjugate Variables* (P or Π^{μ})	C	$\mathring{T}^{\mu\nu}$
Linear Chain	$\begin{cases} \rho\dot{\psi} = P_0 \\ -Y\psi_{,x} = P_1 \end{cases}$	$\frac{1}{2}\left(\frac{1}{\rho}P_0^2 - \frac{1}{Y}P_1^2\right)$	—
Scalar Field	$\phi_{,\mu}$	$\frac{1}{2}\Pi^{\mu}\Pi_{\mu} + m^2\phi^2$	$\Pi^{\mu}\Pi^{\nu} - \frac{1}{2}g^{\mu\nu}(\Pi^{\sigma}\Pi_{\sigma} - m^2\phi^2)$
Complex Scalar Field	$-\frac{1}{2}\phi^{*}{}_{,\mu}$	$\frac{1}{2}\Pi^{\mu}\Pi^{*}{}_{\mu} + m^2\phi\phi^{*}$	$\Pi^{\mu}\Pi^{*\nu} + \Pi^{*\mu}\Pi^{\nu} - g^{\mu\nu}(\Pi^{\sigma}\Pi^{*}{}_{\sigma} - m^2\phi\phi^{*})$
Free Electromagnetic Field (L_{I})	$\Pi^{\alpha\mu} = F^{\alpha\mu}$	$F^{\mu\nu}A^{\mu}{}_{,\nu} + \frac{1}{4}F^{\sigma\varrho}F_{\sigma\varrho}$	$F^{\sigma\mu}A_{\sigma,}{}^{\nu} + \frac{1}{4}g^{\mu\nu}F^{\sigma\varrho}F_{\sigma\varrho}$
Free Electromagnetic Field (L_{III})	$\Pi_{\alpha}{}^{\mu} = -A_{\alpha,}{}^{\mu}$	$-\frac{3}{2}\Pi_{\mu}{}^{\nu}\Pi^{\mu}{}_{\nu}$	$-\Pi_{\alpha}{}^{\mu}\Pi^{\alpha\nu} - g^{\mu\nu}(\frac{1}{2}\Pi_{\alpha}{}^{\sigma}\Pi^{\alpha}{}_{\sigma})$

If we use the field equations, Eqs. 3.52, the second set of canonical equations can be written as

$$\frac{\partial \mathscr{H}}{\partial \psi^{\alpha}} = -\frac{\partial}{\partial x^{\mu}}(\Pi_{\alpha}^{\mu}),$$

or

$$\dot{\Pi}_{\alpha}^{0} = -\frac{\partial \mathscr{H}}{\partial \psi^{\alpha}} = \frac{\partial}{\partial x^{k}}(\Pi_{\alpha}^{k}), \qquad k = 1, 2, 3. \tag{3.103}$$

Equations 3.102 and 3.103 have the usual form of canonical equations, except for the additional term in Eq. 3.103. The field variables ψ^{α} correspond to "coordinates" and the zero components of the conjugate momenta, Π_{α}^{0}, correspond to "momenta" of the system.

We also find that if the Lagrangian density L does not depend explicitly on time, the Hamiltonian density is constant in time. For

$$\frac{d\mathscr{H}}{dt} = \frac{\partial \mathscr{H}}{\partial t} = \Pi_{\alpha 0}\dot{\psi}^{\alpha,0} + \dot{\Pi}_{\alpha 0}\psi^{\alpha,0} - \frac{\partial L}{\partial \psi^{\alpha}}\dot{\psi}^{\alpha} - \frac{\partial L}{\partial \psi^{\alpha,0}}\dot{\phi}^{\alpha,0} - \frac{\partial L}{\partial \psi^{\alpha,k}}\dot{\psi}^{\alpha,k},$$

and the right-hand side vanishes by virtue of field equations.

It should be remarked that, although the canonical Eqs. 3.101 and 3.103 are not covariant, they are perfectly relativistically invariant. These equations can be written down in every frame. The Hamiltonian density has a definite transformation property under Lorentz transformation; namely it transforms as the 00-component of a tensor. (See also the discussion in Chapter I, Sect. 6.)

The Hamiltonian itself is given by

$$H = \int d^{3}x\mathscr{H} = \int d^{3}x\mathring{T}^{00} = P^{0}. \tag{3.104}$$

6. *Lagrangians Involving Higher-Order Derivatives*[1]

So far we have considered Lagrangians which were functions of field quantities and their *first* derivatives only. There are also physical examples where higher-order field equations, and consequently Lagrangians depending on the higher derivatives of field variables, occur. Again the field quantities may be any spinor or tensor fields.

---

[1] For the material in this section we follow A. O. Barut and G. Mullen, *Ann. Phys.* (N.Y.), *20* (1962), 184 and 203.

We consider in this section the action principle

$$\delta W_{12} \equiv \delta \int_{\sigma_2}^{\sigma_1} L(\psi^\alpha, \psi_\mu^\alpha, \psi_{\mu\nu}^\alpha, \cdots)\, dx, \tag{3.105}$$

where the subscripts $\mu\upsilon$ mean partial derivatives of field variables ψ, $\psi_\mu^\alpha = \partial\psi^\alpha/\partial x^\mu$, $\psi_{\mu(n)}^\alpha \equiv \partial\psi^\alpha/\partial x^\mu \partial x^\nu \partial x^\lambda \cdots$ (n times) $\cdots$ etc. Let N be the highest derivative occurring in L. We introduce the intrinsic variations of field quantities and their derivatives which do not involve any coordinate transformations: $\delta_0\psi^\alpha(x) = \psi'^\alpha(x) - \psi^\alpha(x)$ and similarly $\delta_0\psi_\mu^\alpha$, $\delta_0\psi_{\mu(n)}^\alpha$, such that

$$\delta_0\psi_{\mu(n)}^\alpha = \partial_{\mu(n)}\delta_0\psi^\alpha, \tag{3.106}$$

where

$$\partial_\mu(n) = \partial_\mu\partial_\nu\partial_\lambda \cdots (n \text{ times}). \tag{3.107}$$

The intrinsic variation of the Lagrangian density is given by

$$\delta_0 L = \sum_{n=0}^{N} \frac{\partial L}{\partial\psi_{\mu(n)}^\alpha}\, \delta_0\psi_{\mu(n)}^\alpha \,. \tag{3.108}$$

Let us introduce the symmetric function

$$L_\alpha^{\mu(n)} \equiv \frac{1}{n!} \sum_{\substack{\text{all permutations}\\ \text{of indices } \mu(n)}} \frac{\partial L}{\partial\psi_{\mu(n)}^\alpha}. \tag{3.109}$$

Then

$$\begin{aligned} L_\alpha^{\mu(n)}\, \partial_{\mu(n)}\delta_0\psi^\alpha = (-1)^n\, \partial_{\mu(n)}L_\alpha^{\mu(n)}\delta_0\psi^\alpha + \partial_\mu[L_\alpha^{\mu(n)}\, \partial_{\mu(n-1)}\delta_0\psi^\alpha \\ - \partial_\mu L_\alpha^{\mu(n)}\, \partial_{\mu(n-1)}\delta_0\psi^\alpha + \cdots + (-1)^{n-1}\, \partial_{\mu(n-1)}L_\alpha^{\mu(n)}\delta_0\psi^\alpha]. \end{aligned} \tag{3.110}$$

With this, $\delta_0 L$ can now be written as follows:

$$\delta_0 L = \sum_{n=0}^{N} (-1)^n\, \partial_{\mu(n)}L_\alpha^{\mu(n)}\delta_0\psi^\alpha + \partial_\mu\left[\sum_{n=0}^{N-1} \Pi_\alpha^{\mu\nu(n)}\, \partial_{\nu(n)}\delta_0\psi^\alpha\right], \tag{3.111}$$

where we have defined

$$\Pi_\alpha^{\mu\nu(n)} = \sum_{m=0}^{N-(n+1)} (-1)^m\, \partial_{\lambda(m)}L_\alpha^{\mu\nu(n)\lambda(m)}. \tag{3.112}$$

Here $\mu\nu(n)$ denotes a single index μ and a set of n indices ν. Now, as usual, we separate the action principle into a part involving an intrinsic variation of L due to the variation of field components and a part involving the change of the region of integration due to the displacement δx^μ:

$$\int_{\sigma_2}^{\sigma_1} \delta_0 L\, dx + \left(\int_{\sigma_1} - \int_{\sigma_2}\right) d\sigma_\mu\, \delta x^\mu L = 0.$$

If we insert $\delta_0 L$ into this equation and make use of Gauss's theorem (see Sect. 3), we obtain the Euler-Lagrange equations

$$\sum_{n=0}^{N} (-1)^n\, \partial_{\mu(n)} L_\alpha^{\mu(n)} = 0, \tag{3.113}$$

or, written out explicitly,

$$\frac{\partial L}{\partial \psi^\alpha} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial \psi^\alpha_\mu}\right) + \frac{\partial^2}{\partial x^\mu \partial x^\nu}\left(\frac{\partial L}{\partial \psi^\alpha_{\mu\nu}}\right) - \cdots = 0.$$

We also obtain the generator for infinitesimal transformations

$$F(\sigma) = \int_\sigma d\sigma_\mu \left[L\delta x^\mu + \sum_{n=0}^{N-1} \Pi_\alpha^{\mu\nu(n)}\, \partial_{\nu(n)}\delta_0 \psi^\alpha\right], \tag{3.114}$$

so that

$$\delta W_{12} = F(\sigma_1) - F(\sigma_2). \tag{3.115}$$

It can be verified easily that these formulae reduce to the ones discussed in Sect. 3 for the case where L depends on the first derivatives only.

We shall now develop a canonical formalism and obtain expressions for energy, momentum, charge, etc. For this purpose, we note first that, given $\delta_0\psi^\alpha$, the variations $\delta_0\partial_{\mu(n)}\psi^\alpha$ cannot be arbitrary. We must express the generator $F(\sigma)$ in terms of the independent variations of the field quantities. Now the variations $\delta_0\partial_{\mu(n)}\psi^\alpha$ that are normal to the space-like surface are arbitrary. Therefore, we separate the derivatives ∂_μ into a normal directional derivative $\hat\partial$ and a tangential derivative $\underline{\partial}_\mu$ as follows:

$$\hat\partial \equiv n_\mu \partial^\mu, \qquad \underline{\partial} \equiv \partial_\mu - n_\mu \hat\partial, \tag{3.116}$$

where n^μ is the unit time-like vector, $n^\mu n_\mu = 1$. We have then

$$n^\mu \underline{\partial}_\mu = n^\mu \partial_\mu - n^\mu n_\mu \hat\partial = 0;$$

i.e., n^μ is perpendicular to $\underline{\partial}_\mu$; in a special frame n^μ: (1000) $\hat\partial = \partial_0$. We also put

$$\hat\partial^{(n)} = \hat\partial\hat\partial \cdots (n \text{ times}), \qquad n_{\mu(n)} = n_\mu n_\nu \cdots (n \text{ times}),$$

We restrict ourselves to surfaces such that the variations approach zero sufficiently rapidly, at spatial infinity, so that

$$\int d\sigma_\mu\, \underline{\partial}_\nu f^{\mu\nu} = 0;$$

this integral can be transformed into a surface integral at infinity. (See footnote at the beginning of this section.) We can then write the second term of $F(\sigma)$ as follows:

$$\int d\sigma_\mu \sum_{n=0}^{N-1} \Pi_\alpha^{\mu\nu(n)}\, \partial_{\nu(n)}\delta_0\psi^\alpha = \int d\sigma_\mu \sum_{n=0}^{N-1} \Pi_\alpha^{\mu(n)}(\underline{\partial}_\nu + n_\nu\hat{\partial})_{(n)}\delta_0\psi^\alpha,$$

and, after some calculation,

$$= \int d\sigma_\mu \sum_{k=0}^{N-1} \Pi_\alpha^{(k)\mu}\delta_0(\hat{\partial}^{(k)}\psi^\alpha),$$

where we have introduced

$$\Pi_\alpha^{(k)\mu} = \sum_{n=k}^{N-1} (-1)^{n+k}\binom{n}{k} n_\nu^{(k)}\partial_{\nu(n-k)}\Pi_\alpha^{\mu\nu(n)}, \tag{3.118}$$

with

$$\binom{n}{m} = \frac{n!}{(n-m)!m!}.$$

Then the generator becomes

$$F(\sigma) = \int_\sigma d\sigma_\mu\left[\sum_{k=0}^{N-1} \Pi^{(k)\mu}\delta_0(\hat{\partial}^{(n)}\psi^\alpha) - L\delta x^\mu\right]. \tag{3.119}$$

We now introduce the total variation $\bar{\delta}\psi^\alpha$, defined by

$$\bar{\delta}\psi^\alpha = \delta_0\psi^\alpha + \psi^\alpha(x) - \psi'^\alpha(x) = \delta_0\psi^\alpha + \psi_\mu^\alpha(x)\delta x^\mu - \tfrac{1}{2}\varepsilon^{\mu\nu}S_{\mu\nu}^{\alpha\beta}\psi^\beta(x). \tag{3.120}$$

For higher derivatives we obtain

$$\bar{\delta}\psi_{\mu(n)}^\alpha = \delta_0\psi_{\mu(n)}^\alpha + \psi_{\nu\mu(n)}^\alpha\delta x^\nu - \tfrac{1}{2}\varepsilon^{\lambda\sigma}S_{\lambda\sigma}^{\alpha\beta}\psi_{\mu(n)}^\beta. \tag{3.121}$$

When this expression is multiplied by $n_{\mu(n)}$, we get

$$\bar{\delta}(\hat{\partial}^{(n)}\psi^\alpha) = \delta_0(\hat{\partial}^{(n)}\psi^\alpha) + \partial_\nu(\hat{\partial}^{(n)}\partial^\alpha)\delta x^\nu - \tfrac{1}{2}\varepsilon^{\lambda\sigma}S_{\lambda\sigma}^{\alpha\beta}\hat{\partial}^{(n)}\psi^\alpha.$$

With the help of these expressions we now rewrite $F(\sigma)$ in terms of the total variations:

$$\begin{aligned} F(\sigma) = \int_\sigma d\sigma_\mu\Bigg\{\sum_{k=0}^{N-1} \big[&\Pi^{(k)\mu}\bar{\delta}(\hat{\partial}^{(k)}\psi^\alpha) - \Pi^{(k)\mu}\partial_\sigma(\hat{\partial}^{(k)}\psi^\alpha)\delta x^\sigma \\ &+ \tfrac{1}{2}\varepsilon^{\lambda\sigma}\Pi^{(k)\mu}S_{\lambda\sigma}^{\alpha\beta}\hat{\partial}^{(k)}\psi^\alpha\big] - L\delta x^\mu\Bigg\}. \end{aligned} \tag{3.122}$$

Analogous to the first-order case of Sect. 4, we define the quantities

$$f_{(k)}^{\mu\lambda\nu} \equiv \tfrac{1}{2}[\Pi^{(k)\mu}S^{\lambda\nu}\hat{\partial}^{(k)}\psi^\alpha + \Pi^{(k)\nu}S^{\lambda\mu}\hat{\partial}^{(k)}\psi^\alpha + \Pi^{(k)\lambda}S^{\nu\mu}\hat{\partial}^{(k)}\psi^\alpha], \tag{3.123}$$

which are antisymmetric; $f_{(k)}^{\mu\lambda\nu} = -f_{(k)}^{\lambda\mu\nu}$. Because of this property and the fact that

$$\int d\sigma_\mu\, \partial_\lambda f^{\mu\lambda\nu} = \int d\sigma_\lambda\, \partial_\mu f^{\mu\lambda\nu},$$

we can write finally $F(\sigma)$ in the form

$$F(\sigma) = \int_\sigma d\sigma_\mu \left\{ \sum_{k=0}^{N-1} \Pi_\alpha^{(k)\mu} \bar{\delta}(\hat{\partial}^{(k)}\psi^\alpha) - T^{\mu\nu}\delta x_\nu \right\}, \tag{3.124}$$

where

$$T^{\mu\nu} = \sum_{k=0}^{N-1} \{\Pi_\alpha^{(k)\mu} \partial^\nu (\hat{\partial}^{(k)}\psi^\alpha) - \partial_\lambda f_{(k)}^{\mu\lambda\nu}\} - g^{\mu\nu}L. \tag{3.125}$$

Now, as before, we arrive at the conservation laws by considering coordinate transformations

$$\delta x^\mu = a^\mu + \varepsilon^{\mu\nu} x_\nu, \qquad \text{and} \qquad \bar{\delta}(\hat{\partial}^{(k)}\psi^\alpha) = 0.$$

The corresponding generator is

$$-F_{\delta x}(\sigma) = a^\mu P_\mu + \tfrac{1}{2}\varepsilon^{\mu\nu} J_{\mu\nu}, \tag{3.126}$$

where the energy-momentum vector and the angular momentum are given by

$$P^\nu = \int_\sigma d\sigma_\mu T^{\mu\nu}, \qquad J^{\mu\nu} = \int_\sigma d\sigma_\lambda M^{\lambda\mu\nu}, \tag{3.127}$$

with

$$M^{\lambda\mu\nu} = -M^{\lambda\nu\mu} \equiv T^{\lambda\mu} x^\nu - T^{\lambda\nu} x^\mu. \tag{3.128}$$

If the action integral is invariant under these transformations, we see that P^ν and $J^{\mu\nu}$ are independent of the space-like surface σ; hence they are constants of motion. The differential conservation law is

$$\partial_\mu T^{\mu\nu} = \partial_\lambda M^{\lambda\mu\nu} = 0, \tag{3.129}$$

which in turn implies that $T^{\mu\nu}$ is symmetric.

As an example of intrinsic field transformation, consider the gauge transformation

$$\psi'^\alpha = \psi^\alpha\, e^{\pm i\varepsilon\lambda\delta\lambda}, \qquad \text{and} \qquad \delta x_\mu = 0.$$

Because now $\bar{\delta}({}^{(k)}\hat{\partial}\psi^\alpha) = i\varepsilon_\alpha \lambda\delta\lambda\hat{\partial}^{(k)}\psi^\alpha$, we have

$$F_{\delta\lambda}(\sigma) = \int_\sigma d\sigma_\mu \left[\left\{ \sum_{k=0}^{N-1} \Pi_\alpha^{(k)\mu} (i\varepsilon^\alpha \lambda\delta\lambda\hat{\partial}^{(k)}\psi^\alpha) \right\} \right] \equiv \delta\lambda Q(\sigma), \tag{3.130}$$

where

$$Q(\sigma) = \int d\sigma_\mu j^\mu, \qquad j^\mu = i\varepsilon^\alpha{}_\lambda \sum_{k=0}^{N-1} \Pi_\alpha^{(k)\mu} \hat{\partial}^{(k)}\psi^\alpha \tag{3.131}$$

As usual, we can interpret Q as the total "charge" and j^μ the corresponding current density.

7. *Further Developments and Exercises*

1. Show that Faraday's law expresses a relativity principle for a moving circuit in a fixed magnetic field, on the one hand, and a fixed circuit in a moving magnetic field, on the other.

2. Prove Eqs. 3.17 and 3.20.

3. What is the time dependence of a spatially constant magnetic field?

4. Given that **E** is perpendicular to **B** and $E < B$, find a Lorentz frame in which $\mathbf{E} = 0$.

5. Calculate the field of a uniformly moving charged particle by using a Lorentz transformation. (In Part II a different method will be used for the same problem: see Lienard-Wiechert potential.)

6. Using Eq. 3.42 find the transformation property of the dielectric tensor ε_{ij} and magnetic permeability tensor μ_{ij} of material media under Lorentz transformations.

7. Discuss the derivation of Maxwell's equation in material media from the Maxwell-Lorentz equations by a statistical averaging process.[1]

8. What is the meaning of gauge dependence of the energy of a charged particle as given by Eq. 4.158?

9. For any tensor $A^{\mu\sigma\rho}$, prove the following identity:

$$\begin{aligned} A^{\mu\sigma\rho}{}_{,\mu} - A^{\mu\rho\sigma}{}_{,\mu} &= \tfrac{1}{2}[x^\rho(A^{\mu\sigma\nu} - A^{\sigma\nu\mu} - A^{\nu\mu\sigma} - A^{\nu\sigma\mu} - A^{\sigma\mu\nu} - A^{\mu\nu\sigma})_{,\mu} \\ &\quad + \{\sigma \leftrightarrow \rho\}] = (x^\rho A^{\mu\sigma})_{,\mu} - (x^\sigma A^{\mu\rho})_{,\mu} \end{aligned} \tag{3.132}$$

Solution. First calculate the following three divergences with respect to first, second and third indices in the form:

$$\begin{aligned} A^{\mu\sigma\rho}{}_{,\mu} &= A^{\mu\nu\rho}{}_{,\mu} g^\sigma_\nu = \frac{\partial x^\sigma}{\partial x^\nu} A^{\mu\nu\rho}{}_{,\mu} \\ &= (x^\sigma A^{\mu\nu\rho}{}_{,\mu})_{,\nu} - x^\sigma A^{\mu\nu\rho}{}_{,\mu\nu} \\ &= (x^\sigma A^{\mu\nu\rho}{}_{,\mu})_{,\nu} - (x^\sigma A^{\mu\nu\rho}{}_{,\nu})_{,\mu} + (x^\sigma_{,\mu} A^{\mu\nu\rho}{}_{,\nu}) \\ &= [x^\sigma(A^{\nu\mu\rho} - A^{\mu\nu\rho})_{,\nu}]_{,\mu} + A^{\sigma\nu\rho}_{,\nu}, \\ A^{\sigma\nu\rho}{}_{,\nu} &= (x^\rho A^{\sigma\nu\mu}{}_{,\nu})_{,\mu} - x^\rho A^{\sigma\nu\mu}{}_{,\nu\mu} \\ &= (x^\rho A^{\sigma\nu\mu}{}_{,\nu})_{,\mu} - (x^\rho A^{\sigma\nu\mu}{}_{,\mu})_{,\nu} + x^\rho_{,\nu} A^{\sigma\nu\mu}{}_{,\mu} \\ &= [x^\rho(A^{\sigma\nu\mu} - A^{\sigma\mu\nu})_{,\nu}]_{,\mu} + A^{\sigma\rho\mu}{}_{,\mu}, \end{aligned}$$

---

[1] J. van Vleck, *Theory of Electric and Magnetic Susceptibilities* (Oxford, 1939).
P. Mazur and B. R. A. Nijboer, "Derivation of Maxwell's equations from electron theory," *Physica, 19* (1953), 971.

and

$$A^{\sigma\rho\mu}{}_{,\mu} = (x^\sigma A^{\nu\rho\mu}{}_{,\mu})_{,\nu} - x^\sigma A^{\nu\rho\mu}{}_{,\mu\nu}$$

$$= (x^\sigma A^{\mu\rho\nu}{}_{,\nu})_{,\mu} - (x^\sigma A^{\nu\rho\mu}{}_{,\nu})_{,\mu} + x^\sigma_{,\mu} A^{\nu\rho\mu}{}_{,\nu}$$

$$= [x^\sigma(A^{\mu\rho\nu} - A^{\nu\rho\mu})_{,\nu}]_{,\mu} + A^{\nu\rho\sigma}{}_{,\nu}.$$

Adding these three divergences, we obtain

$$A^{\mu\sigma\rho}{}_{,\mu} = [x^\sigma(A^{\nu\mu\rho} - A^{\mu\nu\rho} + A^{\mu\rho\nu} - A^{\nu\rho\mu})_{,\nu}]_{,\mu} + [x^\rho(A^{\sigma\nu\mu} - A^{\sigma\mu\nu})_{,\nu}]_{,\mu}$$
$$+ A^{\nu\rho\sigma}{}_{,\nu},$$

and

$$A^{\mu\rho\sigma}{}_{,\mu} = (\sigma \leftrightarrow \rho).$$

The difference of the last two equations gives the desired result, Eq. 3.132.

10. Derive the symmetric energy-momentum tensor $T^{\mu\nu}$ from the variation principle involving change of boundary, Eq. 3.57′ and the identity, Eq. 3.132.

Solution. For an infinitesimal homogeneous Lorentz transformation,

$$\delta'\psi^\alpha = \Lambda^\alpha_\beta \psi^\beta(x) = \tfrac{1}{2} S^{\alpha\sigma\rho}_\beta \omega_{\sigma\rho} \psi^\beta.$$

If we insert this into Eq. 3.57′, we obtain

$$\partial_\mu[\tfrac{1}{2} L^\mu_\alpha S^{\alpha\sigma\rho}_\beta \omega_{\sigma\rho} \psi^\beta + (g^{\mu\sigma} L - \psi^{\alpha,\sigma} L^\mu_\alpha)\omega_{\sigma\rho} x^\rho] = 0\,;$$

or, with Eqs. 3.63 and 3.67,

$$\partial_\mu[f^{\mu\sigma\rho}\omega_{\sigma\rho} + \mathring{T}^{\mu\sigma}\omega_{\sigma\rho} x^\rho] = 0\,;$$

or

$$\partial_\mu[\tfrac{1}{2}\omega_{\sigma\rho}(f^{\mu\sigma\rho} - f^{\mu\rho\sigma} + x^\rho \mathring{T}^{\mu\sigma} - x^\sigma \mathring{T}^{\mu\rho})] = 0\,;$$

or, by Eq. 3.91,

$$\partial_\mu \mathring{M}^{\mu\sigma\rho} = \partial_\mu[f^{\mu\sigma\rho} - f^{\mu\rho\sigma} + x^\rho \mathring{T}^{\mu\sigma} - x^\sigma \mathring{T}^{\mu\rho}] = 0.$$

We now apply the identity proved in the previous exercise to $(f^{\mu\sigma\rho} - f^{\mu\rho\sigma})_{,\mu}$ to obtain

$$\partial_\mu[x^\rho(\mathring{T}^{\mu\sigma} + T^{\mu\sigma}) - x^\sigma(\mathring{T}^{\mu\rho} + T^{\mu\sigma})] = 0,$$

or

$$(x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho})_{,\mu} = 0,$$

where the symmetric energy-momentum tensor is given by

$$T^{\mu\nu} = \mathring{T}^{\mu\nu} + \tfrac{1}{2}(f^{\alpha\mu\nu} + f^{\mu\nu\alpha} + f^{\nu\alpha\mu} - f^{\mu\alpha\nu} - f^{\alpha\nu\mu} - f^{\nu\mu\alpha})_{,\alpha},$$

which, by virtue of the antisymmetry of $f^{\mu\sigma\rho}$, Eq. 3.75 reduces to

$$T^{\mu\nu} = \overset{\circ}{T}{}^{\mu\nu} + (f^{\alpha\mu\nu} + f^{\mu\nu\alpha} + f^{\nu\alpha\mu})_{,\alpha}. \tag{3.133}$$

11. Prove the algebraic steps leading to Eq. 3.46, the properties of the β matrices, their reduction and finally the two-component forms of Eq. 3.46.

12. Prove Eqs. 3.69 and evaluate $S^{\alpha\mu\nu}_{\beta}$ for two-component spinors. [In the transformation equation

$$\xi'^{\alpha}(x') = A^{\alpha}_{\beta}(L)\xi^{\beta}(x),$$

one has to express $A^{\alpha}_{\beta}(L)$ for an infinitesimal Lorentz transformation $L^{\mu}_{\nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\nu}$, and write $A^{\alpha}_{\beta}(L) = \delta^{\alpha}_{\beta} + \frac{1}{2}S^{\alpha\beta\mu\nu}\omega_{\mu\nu}$, which gives S.]

13. Calculate the angular-momentum tensors $J^{\mu\nu}$ from the energy-momentum tensor $T^{\mu\nu}$ for the fields given in Table II. Discuss the physical meaning of the angular momentum of the electromagnetic field in detail.

14. Calculate the conserved quantities, as in Table II, for the Dirac field. The Lagrangians are given in Table I.

15. *Conformal Invariance.* It has been known for a long time that Maxwell's equations are invariant with respect to a larger class of transformations than the Lorentz transformations, namely the so-called group of *conformal* transformations. [H. Bateman, *Proc. Lond. Math. Soc.*, 8, (1910) 228; 8, (1910) 469; E. Cunningham, *Proc. Lond. Math. Soc.*, 10, (1910) 77.] The general conformal transformations have 15 parameters (a 15-parameter Lie group). They include restricted inhomogeneous Lorentz transformations (10 parameters), a scale transformation $x'_{\mu} = \lambda x_{\mu}$ (1 parameter) and the nonlinear transformations

$$x'^{\mu} = \frac{1}{N}(x^{\mu} - x^2 a^{\mu});$$

$$N = 1 - 2a^{\mu}x_{\mu} + a^2 x^2,$$

where a^{μ} is a constant vector. Because $x'^{\mu}x'_{\mu} = \frac{1}{N}\, x^{\mu}x_{\mu}$, this is still a transformation from flat space to flat space. [A generalization to arbitrary metric $g^{\mu\nu}(x)$ can be made, which is not considered here.] For a special choice $a^{\mu}:(-g/2, 0, 0, 0)$ and in the nonrelativistic limit $c \leftrightarrow \infty$, the above nonlinear transformation describes the passage to a uniformly accelerated frame of reference: $t' = t/(1 - \frac{1}{4}g^2t^2)$ and $z' = -\frac{1}{2}gt^2/(1 - \frac{1}{4}g^2t^2)$.

Show the invariance of Maxwell's equations under this larger class of transformations.

16. *Derivation of Spin Precession from Field Theory.* Consider a spinor field ψ interacting with the electromagnetic field. Using the formalism of

Section 4, in particular equation 3.77, evaluate the angular momentum $J_{\mu\nu}$. The total angular momentum has two parts, one coming from the canonical field tensor $\overset{\circ}{T}{}^{\mu\nu}$, the other from the spin part. Correspondingly we may introduce two dual vectors t_μ and W_μ by

$$t_\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\lambda\rho}p^\nu J^{\lambda\rho}_{\text{orbital}}$$

$$W_\mu = \tfrac{1}{2}\varepsilon_{\mu\nu\lambda\rho}p^\nu J^{\lambda\rho}_{\text{spin}}$$

Derive the equations of the motion for these vectors from the divergence of $T^{\mu\nu}_{,\nu}$. Compare the equation of the motion of the spin vector W_μ with that discussed in Chapter II, Section 4. (See also Chapter I, Problem 14, for the definition of spin from the Lorentz group).

BIBLIOGRAPHY FOR CHAPTER III

Books

1. R. Becker and F. Sauter, *Theorie der Elektrizität, Band II, Elektronentheorie* (Leipzig: Teubner, 1939).
2. N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (New York: Interscience Publishers, 1959), Chapter I.
3. D. Ivanenko and A. Sokolov, *Die Klassische Feldtheorie* (Berlin: Akademie-Verlag, 1953).
4. J. D. Jackson, *Classical Electrodynamics* (New York: John Wiley & Sons, 1962).
5. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Cambridge, Mass.: Addison-Wesley, 1951).
6. W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism* (Cambridge, Mass.: Addison-Wesley, 1955).
7. A. Sommerfeld, *Lectures on Theoretical Physics* (New York: Academic Press, 1950), Vols. 3 and 4.
8. J. Rzewuski, *Field Theory. Part I, Classical Theory* (Warszawa: 1958). (Polish Academy of Sciences—Physical Monograph.)
9. W. Thirring, *Principles of Quantum Electrodynamics* (New York: Academic Press, 1958).
10. M. Mason and W. Weaver, *The Electromagnetic Field* (New York: Dover Publications, 1952).

Articles

1. E. Noether, "Invariante Variations probleme" Göttingen Nachrichten, 211 (1918).
2. W. Pauli, "Relativistic Field Theories of Elementary Particles," *Rev. Mod. Phys.*, *13* (1941), 203.

3. F. Bopp, "Konforminvarianz der Electrodynamik," *Annalen der Physik*, *4*, (1959), 96.
4. R. Utiyama, "Theory of Invariant Variation and the Generalized Canonical Dynamics, *Prog. Theor. Phys.*, Suppl. No. *9* (1959), 19.
5. J. G. Taylor, "Classical Electrodynamics as a Distribution Theory," *Proc. Cambridge Phil. Society*, *52* (1956), 119.
6. H. E. Moses, "A Spinor Representation of Maxwell's Equations," *Nuovo Cimento*, Suppl. *7* (1958), 1; Vachaspati, "Maxwell Equations in the Form of Two-Component Equations, *Proc. Nat. Inst. Sciences India*, *26A* (1960), 359.
7. C. Fronsdal, "On the Theory of Higher Spin Fields," *Nuovo Cimento*, Suppl., *9* (1958), 410.
8. D. W. Sciama, "On a Geometrical Theory of the Electromagnetic Field," *Nuovo Cimento*, *8* (1958), 417.
9. M. Tischer and S. Hess, "Materialgleichungen in Beliebigen Medien," *Annalen der Physik*, *3* (1959), 113.
10. J. Winogradzki, "On the Metrical Energy-momentum Tensor and Noether's Theorem" *Cahier de Physique*, No. 67 (1956).
11. T. S. Chang, "Field Theories with High Derivatives," *Proc. Cambridge Phil. Society*, *44* (1948), 76.

PART II

Interactions of Fields and Particles

IV.

EQUATIONS OF MOTION AND THEIR SOLUTIONS

The description given in the first part of this book of treating the particles and fields separately is essentially an approximate one. As we have noted before, an accelerating particle produces a field itself which changes the external field in which the particle is moving. This new field in turn will effect the motion of the particle. Thus, the theory of interacting fields and particles is intricate and involved, both conceptually and in its mathematical formulation. In this part we shall build up this theory gradually, starting from simple interactions. We give first a theory from the physicist's point of view, using δ functions, etc. The question of mathematical consistency and completeness will be discussed at the end.

The three most important aspects of theory of interactions that we shall emphasize are (1) the problem of radiation and radiation reaction, (2) the equation of motion of particles with radiation reactions, and (3) the consistency, completeness and limitations of the classical theory.

We begin with the Lagrangian form of interactions and the equation of motion.

1. *Interaction of Fields with External "Currents"*

Let $\psi^{\alpha}(x)$ be the set of field variables and $L_0^{(F)}$ the Lagrangian density of the free fields. In the presence of a variable external "current" $j(x)$ which acts as a source for the fields the field equations are modified. The precise way in which the current $j(x)$ acts, therefore its definition, can be given in terms of the Lagrangian formalism. We treat the interaction of the field $\psi^\alpha(x)$ with the current $j(x)$ by an additional term in the Lagrangian density

$$L = L_0^{(F)} + L_1(j(x), \psi^\alpha(x)). \tag{4.1}$$

Later we shall give explicitly $j(x)$ in terms of the particle coordinates. Let it be a given function for the time being.

As an example, we consider the electromagnetic field A_μ. The simplest scalar interaction term with a given current four-vector $j^\mu(x)$ is of the form $A^\mu j_\mu$, and Eq. 4.1 takes the form

$$L = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}A_\mu j^\mu. \tag{4.2}$$

The field equations corresponding to this Lagrangian, as obtained from Eq. 3.52, are

$$F^{\mu\nu}{}_{,\nu} = -\frac{1}{c}j^\mu(x). \tag{4.3}$$

These are precisely the Maxwell equations, Eq. 3.34, so that the interaction written in Eq. 4.2 is the correct one.

The Lagrangian given in Eq. 4.2 is not gauge invariant. However, under the gauge transformations of the potentials A_μ, the sources that produce the field also undergo such a transformation as to produce the continuity equation

$$j^\mu_{,\mu} = 0, \tag{4.4}$$

and as to make the complete Lagrangian gauge invariant, as we shall show later. We note here that under a gauge transformation, $A_\mu \to A_\mu + \Lambda_{,\mu}$, and under the assumption of Eq. 4.4, the Lagrangian Eq. 4.2 gets only a divergence term $-1/c(\Lambda j^\mu)_{,\mu}$.

The total energy-momentum tensor derived from Eq. 4.2 is

$$\Theta^{\mu\nu} = T^{\mu\nu} + \frac{1}{c}(A_\sigma j^\sigma g^{\mu\nu} - A^\nu j^\mu), \tag{4.5}$$

where $T^{\mu\nu}$ is the energy-momentum tensor of the free electromagnetic field, as we have obtained before,

$$T^{\mu\nu} = F^{\mu\sigma}F_\sigma{}^\nu - g^{\mu\nu}L. \tag{4.5$'$}$$

Because the Lagrangian Eq. 4.2 is not invariant under translations, we expect that $\Theta^{\mu\nu}$ does not satisfy an equation of continuity. Indeed, we get from Eq. 4.2, using Eq. 3.94,

$$\Theta^{\mu\nu}{}_{,\mu} = -\left.\frac{\partial L}{\partial X_\nu}\right|_{\text{expl}} = \frac{1}{c}A_\mu j^{\mu,\nu}. \tag{4.6}$$

On the other hand, from Eq. 4.5,

$$\Theta^{\mu\nu}{}_{,\mu} = T^{\mu\nu}{}_{,\mu} + \frac{1}{c}(A_{\sigma}^{;\nu} j^{\sigma} + A_{\sigma} j^{\sigma,\nu} - A^{\nu}_{,\sigma} j^{\sigma} - A^{\nu} j^{\sigma}_{,\sigma}).$$

Comparing this equation with Eq. 4.6 and using Eq. 4.4, we obtain

$$T^{\mu\nu}{}_{,\mu} = \frac{1}{c} F^{\mu\nu} j_{\mu} \equiv k^{\nu}, \tag{4.7}$$

where we have also introduced a new four-vector k^{μ} which we shall call the *force density*. The physical interpretation of Eq. 4.7 is as follows. If the field is "driven" by a current $j(x)$, the energy-momentum tensor no longer satisfies a continuity equation, as in the case of free fields. In other words, the energy and momentum of the field changes as a function of time according to the equation

$$P^{\nu}(\sigma_2) - P^{\nu}(\sigma_1) = \int_R dx k^{\nu}, \tag{4.8}$$

where R is the region between the two space-like surfaces σ_1 and σ_2. Or, in a particular frame of reference,

$$dP^{\nu}/dt = \int d^3x k^{\nu} \equiv K^{\nu}; \tag{4.8$'$}$$

i.e., the rates of change of the total field energy P^0 and total field momentum are given by the space integral of the force density, or the total force K^{ν}, which is, of course, the analogue of the Minkowski equation for a single particle discussed in Chapter II.

It is instructive to derive the fundamental equation, Eq. 4.7, from the definition, Eq. 4.5′, of $T^{\mu\nu}$ and the field equations, Eq. 4.3. Note that in the above derivation we have not used field equations but a property of the Lagrangian that being not invariant under translations. Differentiating Eq. 4.5′, we obtain

$$\begin{aligned} T^{\mu\nu}{}_{,\mu} &= F^{\sigma\mu}{}_{,\mu} F^{\nu}{}_{\sigma} + F^{\sigma\mu} F^{\nu}_{\sigma,\mu} + \partial^{\nu}(\tfrac{1}{4} F^{\sigma\rho} F_{\sigma\rho}) \\ &= -\frac{1}{c} j^{\sigma} F^{\nu}{}_{\sigma} + F^{\sigma\mu} F^{\nu}_{\sigma,\mu} + \tfrac{1}{2} F^{\sigma\rho,\nu} F_{\sigma\rho}. \end{aligned}$$

We now make use of the identity, Eq. 3.33, or

$$\tfrac{1}{2} F_{\sigma\rho} F^{\sigma\rho,\nu} = \tfrac{1}{2} F_{\sigma\rho}(-F^{\rho\nu,\sigma} - F^{\nu\sigma,\rho}),$$

to obtain

$$T^{\mu\nu}{}_{,\mu} = \frac{1}{c} F^{\nu\sigma} j_{\sigma} + \tfrac{1}{2} F_{\sigma\rho} F^{\nu\sigma,\rho} - \tfrac{1}{2} F_{\sigma\rho} F^{\rho\nu,\sigma}.$$

The last two terms cancel and we get again Eq. 4.7. Q.E.D.

Equation 4.7 is not in the form of a conservation law and we have seen that the energy and momentum of the field are not conserved. (The angular momentum is conserved, because the Lagrangian is invariant under homogeneous Lorentz transformations.) If we introduce an energy-momentum tensor $\tau^{\mu\nu}$ for the current; i.e., the matter tensor, by the equation

$$\tau^{\mu\nu}{}_{,\mu} = -k^{\nu} = -\frac{1}{c} F^{\mu\nu} j_{\mu}, \tag{4.9}$$

we can express Eq. 4.7 in the form of an equation of continuity:

$$(T^{\mu\nu} + \tau^{\mu\nu})_{,\mu} = 0. \tag{4.10}$$

Consequently the total energy and momentum of the field *and* the matter are conserved; they are not conserved separately. There is an exchange of energy and momentum between the two systems.

2. *Interaction of Fields with a Particle*

The discussion in the previous section has shown that from the point of view of both field equations and conservation laws it is desirable, or necessary, to specify the nature of the current $j(x)$ explicitly. Now, we consider the system of fields and particles together when there is an energy exchange between them. The current $j(x)$ will be a function of the particle coordinates $z^{\mu}(s)$ and their derivatives. The total Lagrangian can now be written in the following form:

$$L = L_0^{(F)}(\psi^{\alpha}, \psi^{\alpha,\mu}) + L_0^{(P)}(z^{\mu}, \dot{z}^{\mu}) + L_1(\psi^{\alpha}, \psi^{\alpha,\mu}, z^{\mu}, \dot{z}^{\mu}), \tag{4.11}$$

which consists of terms corresponding to free fields and particles plus an interaction term. If we vary L with respect to the particle coordinates, we obtain the equation of motion

$$m_0 c \ddot{z}^{\mu} = K^{\mu}(\psi^{\alpha}), \tag{4.12}$$

where the force now depends on the field variables. Similarly, the variation of L with respect to $\psi^{\alpha}(x)$ gives the field equations

$$\Lambda_{\alpha} = f_{\alpha}(z^{\mu}), \tag{4.13}$$

with an inhomogeneous source term due to the presence of the particles. $\Lambda_{\alpha} = 0$ are the free-field equations. The two coupled equations, Eqs. 4.12 and 4.13, then form a complete framework describing the interactions of fields and particles.

Let us consider a single particle, and let $z^\mu = z^\mu(s)$ be the world line of the particle. The current $j^\mu(x)$, as a function of x, is different from zero only along the world line.

For a charged particle interacting with an electromagnetic field, we have already introduced a current four-vector,

$$j^\mu = \rho_0 c u^\mu,$$

which, as a function of x, can be written now as

$$j^\mu(x) = ec \int_{-\infty}^{+\infty} ds \dot{z}^\mu(s) \delta(x - z(s)), \tag{4.14}$$

with

$$\dot{z}^\mu \dot{z}_\mu = 1. \tag{4.15}$$

First we must verify that, for this particular form of the current vector, the equation of continuity, Eq. 4.4, is satisfied, as it should be, because the charge e of the particle is conserved. Indeed,

$$j^\mu_{,\mu} = ec \int_{-\infty}^{+\infty} ds \dot{z}^\mu(s) \frac{\partial}{\partial x^\mu} [\delta(x - z(s))],$$

or, because

$$\partial f(x - z)/\partial x^\mu = -\partial f(x - z)/\partial z^\mu,$$

and

$$(\partial f/\partial z^\mu)(dz^\mu/ds) = df/ds,$$

$$j^\mu_{,\mu} = -ec \int_{-\infty}^{+\infty} ds \frac{d}{ds} [\delta(x - z(s))] = -ec\delta(x - z(s)) \Big|_{s=-\infty}^{+\infty} = 0. \quad Q.E.D.$$

Now we must use the Lagrangian Eq. 4.2 with j^μ given by Eq. 4.14. Under gauge transformation, the Lagrangian changes by a total divergence term so that the equations of motion are gauge invariant. They are, of course,

$$F^{\mu\nu}{}_{,\nu} = -e \int_{-\infty}^{+\infty} ds \dot{z}^\mu(s) \delta(x - z(s)). \tag{4.16}$$

The fundamental Eq. 4.7 is now

$$\overline{T}^{\mu\nu}{}_{,\mu} = eF^{\mu\nu} \int ds \dot{z}_\mu(s) \delta(x - z(s)) = eF^{\mu\nu}(z) \dot{z}_\mu(s) \big|_{z(s)=x}. \tag{4.17}$$

Thus, the divergence of the energy-momentum tensor of the field at x is essentially the Lorentz force on the particle when the particle is at the point x: $z^\mu(s) = x$. This in turn is equal to mass times acceleration, $m_0 c \ddot{z}^\nu(s)$. Indeed, for a free particle with a Lagrangian

$$L_0^{(P)} = m_0 c[(\dot{z}^\mu(\tau))^2]^{1/2},$$

we have the equation

$$\delta L = \frac{\partial}{\partial \tau}\left(\frac{\partial L}{\partial \dot{z}^{\mu}}\,\delta z^{\mu}\right) = \frac{\partial}{\partial z^{\nu}}\left(\frac{\partial L}{\partial \dot{z}^{\mu}}\,\delta z^{\mu}\right) dz^{\nu}/d\tau$$

and under a translation, $\delta z^{\mu} = a^{\mu} = \text{const.}$, and translational invariance, $\delta L = 0$, the equation of continuity is

$$(m_0 c \dot{z}^{\mu} \dot{z}^{\nu})_{,\nu} = 0.$$

We can, therefore, define an energy-momentum or matter tensor for the single particle by (with correct dimensional factors)

$$\tau^{\mu\nu} \equiv -m_0 c^2 \int ds \dot{z}^{\mu} \dot{z}^{\nu} \delta(x - z(s)), \tag{4.18}$$

with $\tau^{\mu}{}_{\mu} = -m_0 c^2$. We want now to show that the sum of the energy-momentum tensors of the field and the particle is conserved.

From Eq. 4.18 we get

$$\tau^{\mu\nu}{}_{,\mu} = -m_0 c^2 \int ds \dot{z}^{\mu} \dot{z}^{\nu}\, \partial_{\mu} \delta(x - z) = m_0 c^2 \int ds \dot{z}^{\nu} \frac{d}{ds}\,[\delta(x - z)]$$

$$= m_0 c^2 \int ds \frac{d}{ds}\,[\dot{z}^{\nu} \delta(x - z)] - m_0 c^2 \int ds \ddot{z}^{\nu} \delta(x - z);$$

the first term vanishes, hence

$$\tau^{\mu\nu}{}_{,\mu} = -m_0 c^2 \int ds \ddot{z}^{\nu} \delta(x - z(s)). \tag{4.19}$$

The conservation of the total energy-momentum tensor $(T^{\mu\nu} + \tau^{\mu\nu})_{,\mu} = 0$ is now satisfied, in view of the equation of motion for the particle

$$m_0 c \ddot{z}^{\nu} = \frac{e}{c} F^{\nu\mu} \dot{z}_{\mu}. \tag{4.20}$$

The complete action governing both the field and the particle is given by

$$\int L\, dx = \int m_0 c (\dot{z}^{\mu 2})^{1/2}\, d\tau + \frac{e}{c} \int A_{\mu}(x) \dot{z}^{\mu}(\tau) \delta(x - z(\tau))\, dx\, d\tau$$
$$+ \tfrac{1}{4} \int dx F^{\mu\nu} F_{\mu\nu}. \tag{4.21}$$

It can be shown that this action leads to the coupled Eqs. 4.18 and 4.20, just discussed. Furthermore, the conservation law derived from Eq. 4.21 is just

$$(T^{\mu\nu} + \tau^{\mu\nu})_{,\mu} = 0,$$

with $T^{\mu\nu}$ and $\tau^{\mu\nu}$ given by Eq. 4.5′ and Eq. 4.18 respectively.

As a second example, we consider a particle interacting with a complex scalar field $\phi(x)$ whose Lagrangian is given in Table I, Chapter III. The simplest interaction term in the action is

$$\int g\phi(x)\delta(x - z(s))\, ds, \tag{4.22}$$

where g is a coupling constant. The fundamental coupled Eqs. 4.12 and 4.13 become in this case

$$(\Box^2 + m^2)\phi(x) = g\delta(x), \tag{4.23}$$

for the field, and

$$m_0 c\ddot{z}^\mu = g \int ds\phi(x)\, \partial^\mu[\delta(x - z(s))] = K^\mu, \tag{4.23'}$$

or the particle. It is easily verified that the Minkowski force in the last equation satisfies the necessary condition

$$K^\mu \dot{z}_\mu = 0,$$

for

$$K^\mu \dot{z}_\mu = g \int ds\phi(x)\dot{z}_\mu[\delta(x - z(s))] = g\phi(x)\delta(x - z)\Big|_{-\infty}^{+\infty} = 0.$$

It must be noted that the classical aspect of the scalar field, such as the π-meson field, is a much less common experience than the electromagnetic field. Whereas in every-day experience we almost always have a very large number of photons and therefore can describe them by the classical electromagnetic field (in other words, we see their field aspect), we can only deal at present with the particle aspect of the π mesons. It is conceivable that, in spite of their massive character, when we must deal with a very large number of π mesons, we also should be able to describe their behavior classically by a pseudoscalar field.

The form of the Lorentz force Eq. 4.20 is experimentally well established; the correct form of the interaction of a particle with the scalar field is not. The form, Eq. 4.22, above is the simplest possibility, which, however, may not be the correct one. In the case of a static field, $\phi = \phi(\mathbf{r})$, $\phi = 0$, Eq. 4.23 reduces to

$$(\Delta^2 - m^2)\phi(\mathbf{r}) = - g\delta(\mathbf{r}), \tag{4.24}$$

and we can give an interpretation similar to the classical gravitational or electrostatic field. The scalar fields have a source, the nuclear matter with a mesonic charge g, and one can regard the mesonic field as arising from a distribution of nuclear matter in the same way as one regards the gravitational field as arising from a distribution of mass, or the electrostatic field from a distribution of charge. The solution of Eq. 4.24 has the form

$$\phi(\mathbf{r}) = \frac{1}{4\pi}\, ge^{-mr}/r$$

(see Chapter IV), which falls much more rapidly than the gravitational and electrostatic fields which go as $1/r$. This is due, of course, to the mass of mesons and makes it even more difficult to observe the classical aspects of such fields.

3. *Interactions Between Fields*

The description of fields and particles and their interactions discussed in the previous section is quite unsymmetrical. The particles are described by coordinates $z^\mu(s)$, the fields by functions $\psi^\alpha(x)$. Classically, field and particle are two different and distinct concepts; hence their descriptions are not on the same basis. Quantum theory, in particular quantum field theory, attempts a synthesis between these two concepts. Here, both particles and fields are described by functions $\psi^\alpha(x)$, in the sense that one field (or particle) is the source of the other. In other words, there are only fields or particles.

Although the field quantities $\psi^\alpha(x)$ considered in quantum theory of fields are not ordinary functions but operators, we briefly consider in this section the interactions of fields with other fields on a classical level. This treatment can be considered as an approximation in some cases and is intended as an introduction to quantized field theory. Furthermore, we want to show how gauge invariance and the equation of continuity for the current $j^\mu_{,\mu} = 0$ are tied together.

Let the charged particles be described by a scalar field $\phi(x)$ and as the source of the electromagnetic field, and vice versa. The total Lagrangian density is of the form

$$L = L_0(\phi) + L_0(A^\mu) + L_1(\phi, A^\mu), \tag{4.25}$$

where $L_0(\phi)$ and $L_0(A^\mu)$ are the Lagrangian densities of the free scalar and electromagnetic fields given in Table I. We determine L_1 from the requirements of Lorentz invariance and gauge invariance. The simplest possibility is not a term of the form $A^\mu j_\mu$, as one might expect, because this form is not gauge invariant. We proceed as follows. We know that $L_0(A^\mu)$ in Eq. 4.25 is invariant under the gauge transformations

$$A_\mu \to A_\mu + \Lambda_{,\mu}. \tag{4.26}$$

The term $L_0(\phi)$ is invariant under the *gauge transformations of the first kind,*

$$\phi \to \phi e^{-ie\lambda},$$

where λ is a constant. If we now require that L be invariant under Eq. 4.26 and a *gauge transformation of the second kind* for ϕ,

$$\phi \rightarrow \phi e^{-ie\Lambda(x)}, \tag{4.27}$$

simultaneously, $L_1(\phi, A^\mu)$ is determined uniquely. The simplest way to see this is to note that the expression $(i\partial_\mu - eA_\mu)\phi$ is invariant under Eqs. 4.26 and 4.27. We thus replace in $L_0(\phi)$ the differential operator ∂_μ by[1]

$$\partial_\mu \rightarrow (i\partial_\mu - eA_\mu), \tag{4.28}$$

and obtain

$$L = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \tfrac{1}{2}[(i\partial_\mu - eA_\mu)\phi(-i\partial^\mu - eA^\mu)\phi^* - m^2\phi\phi^*]; \tag{4.29}$$

or, separating $L_0(\phi)$ out, we have

$$L = L_0(A^\mu) + L_0(\phi) - j_\mu A^\mu + \tfrac{1}{2}e^2 A_\mu A^\mu \phi\phi^*, \tag{4.29$'$}$$

where

$$j_\mu = -\frac{ie}{2}(\phi^*_{,\mu}\phi - \phi_{,\mu}\phi^*) \tag{4.30}$$

is the current density. The fundamental coupled field equations determined from the Lagrangian, Eq. 4.29, are

$$F^{\mu\nu}{}_{,\nu} = -j^\mu + e^2 A^\mu \phi\phi^* = -\frac{e}{c}[\phi(-i\partial^\mu - eA^\mu)\phi^* + \phi^*(i\partial^\mu - eA^\mu)\phi], \tag{4.31}$$

and

$$[(i\partial_\mu - eA_\mu)^2 - m^2]\phi(x) = 0, \tag{4.32}$$

in which one field acts as the source of the other field in a not very obvious way.

The Lagrangian, Eq. 4.29, is invariant under the inhomogeneous Lorentz transformations which leads to a total energy-momentum tensor that satisfies

$$\tau^{\mu\nu}{}_{,\mu} = 0.$$

The energy momentum of both fields together is conserved, as there is exchange of energy and momentum between the fields.

As a second example, we describe the charged particle by a four-component spinor field $\psi(x)$ satisfying the Dirac equation. The Lagrangian is given in Table I, Chapter III. The interaction term in the Lagrangian will be obtained again by the requirement of Lorentz and gauge invariance, Eq. 4.26, and

$$\psi \rightarrow \psi e^{-ie\Lambda(x)}. \tag{4.33}$$

---

[1] $c = 1$ in this section.

If we replace ∂_μ in the Lagrangian for the Dirac field by the gauge-invariant derivative $(i\partial_\mu - eA_\mu)$ according to Eq. 4.28, we obtain

$$L = L_0(A^\mu) + L_0(\psi) - eA_\mu(x)\bar{\psi}\gamma^\mu\psi, \tag{4.34}$$

in which L_1 is again of the simple form $A_\mu j^\mu$, because, for a Dirac field,

$$j^\mu = e\bar{\psi}\gamma^\mu\psi. \tag{4.35}$$

The coupled field equations are now

$$[\gamma^\mu(i\partial_\mu - eA_\mu) + m]\psi = 0, \tag{4.36}$$

and

$$F^{\mu\nu}{}_{,\nu} = -j^\mu = -\, e\bar{\psi}\gamma^\mu\psi. \tag{4.37}$$

In these two examples we have determined the form of interaction by the requirement of gauge invariance of the second kind. We arrive at a fundamental idea if we present the argument as follows. Suppose we have a matter field with a current density of the form of Eq. 4.30 or Eq. 4.35, which satisfies the conservation law, $j^\mu_{,\mu} = 0$, a so-called *conserved current*. We may start with this assumption of charge conjugation. Thus, the matter field is invariant under a gauge transformation of the form of Eq. 4.27 or Eq. 4.33. But the Lagrangian of the matter field alone is not invariant under such a transformation, unless we introduce a vector field $A_\mu(x)$ satisfying the gauge transformation, Eq. 4.26, to counteract the variation of the matter field. Thus, from this point of view, a conservation law, here charge conservation, forces us to introduce a vector field through which the particles interact. The charge we are discussing here is purely the electromagnetic charge. This idea is susceptible to generalization, as will be shown below in the discussion of the so-called Yang-Mills field.

**Vector Fields.*[1] The electromagnetic field $A_\mu(x)$ is a vector field with zero mass which introduces an additional gauge condition. It is important in the theory of elementary particles to consider vector fields with nonzero mass. In this case there are no subsidiary gauge conditions.

Let us denote the vector field by $\phi_\mu(x)$ and its mass by m, and let us investigate directly the electromagnetic interactions of this vector field with the electromagnetic field $A_\mu(x)$. As before, only the so-called gauge-independent derivatives of the vector field,

$$D_\nu\phi_\mu(x) = (i\partial_\nu - eA_\nu)\phi_\mu(x) \tag{4.38}$$

should enter into the Lagrangian; this guarantees gauge invariance with respect to the electromagnetic field.

[1] Sections with * may be omitted at first reading.

We assume a Lagrangian that is bilinear in ϕ_μ, ϕ^*_μ and its gauge-invariant derivatives. The most general such Lagrangian is of the form

$$L = A^{\mu\nu}_{\sigma\rho}(D^*_\mu\phi^{*\sigma})(D_\nu\phi^\rho) + m^2\phi^{*\mu}\phi_\mu, \tag{4.39}$$

where $A^{\mu\nu}_{\sigma\rho}$ is an invariant or isotropic tensor, one which has the same form in every coordinate frame. We can determine the most general form of invariant tensor of rank four by a multiple application of the definition of Lorentz transformations (Chapter I); i.e., $\Lambda^\mu{}_\nu g^{\nu\sigma}\Lambda_\sigma{}^\rho = g^{\mu\rho}$, and obtain

$$A^{\mu\nu}_{\sigma\rho} = \alpha g^{\mu\nu}g_{\sigma\rho} + \beta g^\mu_\sigma g^\nu_\rho + \gamma g^\mu_\rho g^\nu_\sigma, \tag{4.40}$$

where α, β, γ are arbitrary scalar functions; in our case these must be arbitrary constants in order to avoid the introduction of new fields.

The above Lagrangian leads immediately to the field equations

$$A^{\mu\nu}_{\sigma\rho}D_\mu D_\nu\phi^\rho = m^2\phi_9; \tag{4.41}$$

or, explicitly,

$$\alpha D^\mu D_\mu\phi_\sigma + (\beta+\gamma)D_\sigma(D_\nu\phi^\nu) - e\beta\phi^\rho F_{\rho\sigma} = m^2\phi_\sigma, \tag{4.41$'$}$$

where $F^{\mu\nu}$ is the electromagnetic field, and we have used the identity

$$D_\mu D_\nu - D_\nu D_\mu = ieF_{\mu\nu}.$$

It is interesting that, besides m and e, three new constants α, β, γ enter into the theory which must correspond to some properties of these vector particles. To see this, we go back to free particles, where the equation of motion is now

$$\alpha\partial^\mu\partial_\mu\phi_\sigma + (\beta+\gamma)\partial_\sigma(\partial_\nu\phi^\nu) = m^2\phi_\sigma. \tag{4.42}$$

The solutions of this equation can be interpreted best if we make the ansatz

$$\phi_\mu = \psi_\mu + \partial_\mu\phi,$$

where ψ_μ is a divergenceless vector field, $\partial_\mu\psi^\mu = 0$ and ϕ is a scalar field. Then the field equations split into two parts

$$\begin{aligned}(\alpha\partial^\mu\partial_\mu - m^2)\psi_\nu &= 0,\\ (\alpha+\beta+\gamma)\partial^\mu\partial_\mu\phi &= m^2\phi.\end{aligned} \tag{4.43}$$

The first equation represents a vector particle of mass m, the second a scalar particle of mass $m/(\alpha+\beta+\gamma)^{1/2}$. It is remarkable that the starting Lagrangian, which was written down by relativistic and gauge invariance conditions only, contains these two possibilities.

The appearance of vector and scalar fields together is perhaps of some deep significance. But customarily one restricts the theory to that of the vector field by choosing

$$\alpha + \beta + \gamma = 0.$$

One of the constants, say α, can be taken to be 1 (this is a change of overall factor in the Lagrangian), then we have $1 + \beta + \gamma = 0$ or $\beta = -(1 + \gamma)$ so that the Lagrangian contains one more constant in addition to the mass, γ. This constant must correspond to some intrinsic property of the particle which we will now determine. With $\alpha = 1$, $\beta = -1 - \gamma$, Eq. 4.40 becomes

$$A^{\mu\nu}_{\sigma\rho} = g^{\mu\nu} g_{\sigma\rho} - g^{\mu}_{\sigma} g^{\nu}_{\rho} + \gamma(g^{\mu}_{\rho} g^{\nu}_{\sigma} - g^{\mu}_{\sigma} g^{\nu}_{\rho}).$$

The Lagrangian density is now

$$L = \tfrac{1}{2}(D^{*}_{\mu}\phi^{*}_{\nu} - D^{*}_{\nu}\phi^{*}_{\mu})(D^{\mu}\phi^{\nu} - D^{\nu}\phi^{\mu}) + m^2 \phi^{*\mu}\phi_{\mu}$$
$$+ \gamma[(D^{*}_{\mu}\phi^{*\mu})(D_{\nu}\phi^{\nu}) - (D^{*}_{\nu}\phi^{*\mu})(D_{\mu}\phi^{\nu})].$$

The term in square brackets can be rewritten, and is equal

$$-\frac{ie}{2}\gamma F_{\mu\nu}(\phi^{*\mu}\phi^{\nu} - \phi^{*\nu}\phi^{\mu}),$$

plus two divergence terms, $-\partial_{\nu}(\phi^{*\mu} D_{\mu}\phi^{\nu}) + \partial_{\mu}(\phi^{*\mu} D_{\nu}\phi^{\nu})$, which may be dropped in L (see Chapter III). The equations of motion which follow from this final form of the Lagrangian are

$$D^{\mu}F_{\mu\nu} = m^2 \phi_{\nu} + ie\gamma\phi^{\mu}F_{\mu\nu},$$
$$D_{\mu}\phi_{\nu} - D_{\nu}\phi_{\mu} = F_{\mu\nu}.$$

This equation and the current density, obtained according to the general principles of the previous chapter,

$$j^{\mu} = ie[\phi^{*}_{\nu}(D^{\mu}\phi^{\nu} - D^{\nu}\phi^{\mu}) - (D^{*\mu}\phi^{*\nu} - D^{*\nu}\phi^{*\mu})\phi_{\nu} + \gamma\partial_{\nu}(\phi^{*\mu}\phi^{\nu} - \phi^{*\nu}\phi^{\mu})]$$
$$= ie[(\phi^{*}_{\mu}D^{\mu}\phi^{\nu} - \phi_{\nu}D^{*\mu}\phi^{*\nu}) + (\gamma + 1)\partial_{\nu}(\phi^{*\mu}\phi^{\nu} - \phi^{*\nu}\phi^{\mu}) + \phi^{\mu}(D^{*}_{\nu}\phi^{*\mu} -$$
$$- \phi^{*\mu}(D_{\nu}\phi^{\nu})],$$

allows one, then, to interpret $(1 + \gamma)$ as the magnetic moment of the vector particle.

* *Yang-Mills and Gauge Fields.* We have seen that the conservation of charge for a matter field, or equivalently the equation of continuity $j^{\mu}_{,\mu} = 0$ for the current density, introduced the gauge-invariant electromagnetic field $A_{\mu}(x)$, if the total Lagrangian has to be invariant under gauge transformations of the second kind.

There are other conservation laws satisfied by interactions of elementary particles besides charge conservation and the conservation laws associated with coordinate transformations.

The elementary particles, or fields of a given kind, occur usually in the form of multiplets. That is to say, we have a set of n fields

$$\psi^{\alpha}(x),\ \alpha = 1, 2, \cdots, n,$$

with the same spinor or tensor character. These fields can be considered as vectors in a finite n-dimensional vectors space, the so-called *internal space*. According to our general discussion in the previous chapter, we can consider now variations of the field quantities corresponding to "rotations" in internal space; for example,

$$\psi'_{\alpha}(x) = S_{\alpha}{}^{\beta}\psi_{\beta}(x); \tag{4.44}$$

i.e., the transformed fields ψ'_{α} are linear combinations of the old fields at the same space-time point x. These transformations are generalizations of the simple gauge transformations that we have considered before. In accordance with our general theory, the transformations of Eq. 4.43 will lead to conservation laws that will be then the generalization of charge conservation.

An important example of such a law is found among the charge multiplets of particles occurring in nuclear forces, or so-called strong interactions. Here the charge multiplets of the particles are grouped into singlets, doublets or triplets, and the corresponding fields transform like scalar, spinors or vectors, respectively. Actually, in nature the fields transform according to the representations of the three-dimensional rotation group; this three-dimensional space is called the *isotopic spin space*, in analogy to the ordinary spin. The use of group representations here is similar to the use of representations of the Lorentz group, as discussed in Chapters I and III, but now it is used for the index α of the fields $\psi_{\alpha}(x)$, instead of the coordinates x. The rotation group admits scalar, tensor and spinor representations.

If the total Lagrangian is invariant under the transformations of Eq. 4.44, we will obtain a conservation law; the conserved quantity is the generator of the infinitesimal transformations corresponding to Eq. 4.44; it is called the *isotopic spin*. There is also a corresponding "conserved" current density j^{μ}, satisfying $j^{\mu}_{,\mu} = 0$.

Of course, it is possible to develop a theory with a more general symmetry group in internal space than the three-dimensional rotation group we have assumed. The steps of such a theory are the same as above; we have to use fields that transform according to representations of this more general symmetry group. There are indications that the elementary particles may obey approximately a more general internal symmetry.

Now, if we start from these conservation laws, such as the isotopic spin current, and require the invariance of the total Lagrangian under transformations of the type

$$\delta\psi = g\omega^{\alpha}(x)\sigma_{\alpha}\psi, \tag{4.45}$$

where σ_{α}, $\alpha = 1, 2, 3$ are the Pauli matrices, which are allowed under $j^{\mu}_{,\mu} = 0$, then we are forced to introduce three-vector fields $b^{\alpha}_{\mu}(x)$ coupled to the matter field ψ to counteract the variation of the free-field Lagrangian corresponding to ψ. This vector field, generalizing the electromagnetic field, is called the *Yang-Mills field.*

One can go one step more. There are two other conservation laws known in strong interactions, the conservation of *nucleon number*, which accounts for the stability of matter, and a new type of charge, the so-called *hypercharge* which accounts for some properties of the so-called new *strange particles*. These two conservation laws are associated with simple gauge transformations of the fields. Following the previous discussion they give rise also to two new massive vector fields, one coupled to hypercharge, the other to nucleon number.

To get a classical feeling of these important questions of the quantum theory of elementary particles, the interested reader should consider the Lagrangians, the infinitesimal transformations of the fields, generators and currents. They are straightforward applications of our general formalism developed in Chapter III. (See exercises at the end of this chapter.)

4. *Solutions of Field Equations: Green's Functions*

The discussion and the interpretation of the coupled equations of motion for the field and the particle, derived in Section 2, will make up essentially the rest of the book. This is no easy matter, as the solution of the coupled partial differential equations is quite involved. In this section we discuss the solution of one of the coupled equations and illustrate a powerful method of finding the solutions, namely the *method of Green's functions.*

Let us consider first the case of electrodynamics. Equation 4.3, with the choice of gauge

$$A^{\mu}_{,\mu} = 0, \tag{4.46}$$

reduces to

$$\Box^2 A^\mu = \frac{1}{c} j^\mu. \tag{4.47}$$

We shall find the general solution of Eq. 4.47, and consider the auxiliary condition, Eq. 4.46, subsequently.

If $G(x - x')$ is a solution of the equation

$$\Box^2 G(x - x') = \delta^4(x - x'), \tag{4.48}$$

then the most general solution of Eq. 4.47 is of the form

$$A^\mu(x) = \underset{0}{A}{}^\mu(x) + \frac{1}{c} \int dx' G(x - x') j^\mu(x'), \tag{4.49}$$

where $\underset{0}{A}{}^\mu(x)$ is any solution of the homogeneous equation; i.e.,

$$\Box^2 \underset{0}{A}{}^\mu(x) = 0. \tag{4.50}$$

Clearly $A^\mu(x)$ given by Eq. 4.49 satisfies Eq. 4.47. We must keep in mind from the beginning that the solutions of Eqs. 4.48 and 4.50 are not unique, unless specified by boundary conditions. Additional difficulties will arise due to the δ-function character of $j^\mu(x)$ in the case of point particles.

The Green's function $G(x - x')$ has a simple physical meaning. It represents the field at the point x due to a unit current density at the point x'. In Eq. 4.49 we integrate over all points x' with a weight factor $j^\mu(x)$ to obtain the total field due to currents.

Green's functions can be obtained by taking the Fourier transform of Eq. 4.48. Let

$$G(x - x') = (2\pi)^{-4} \int dk G(k) e^{-ik(x-x')}; \tag{4.51}$$

we also expand the δ function[1]

$$\delta(x - x') = (2\pi)^{-4} \int dk e^{-ik(x-x')}. \tag{4.52}$$

---

[1] From the defining equation for the δ function

$$\int dx' f(x') \delta(x - x') = f(x),$$

and using (4.51) for $f(x)$, we get

$$\int dx' \, dk \, f(k) e^{-ikx'} \, \delta(x - x') = 1,$$

or

$$\int dx' \, e^{ik(x-x')} \, \delta(x - x') = 1;$$

i.e., the Fourier transform of the δ function is one.

Inserting Eqs. 4.51 and 4.52 into Eq. 4.48 and operating with $\Box^2$ under the integral (differentiation and integration have been interchanged), we obtain

$$-\int dk G(k) k^2 e^{-ik(x-x')} = \int dk e^{-ik(x-x')}.$$

Hence the Fourier transform of $G(x - x')$ is given by

$$G(k) = -1/k^2; \tag{4.53}$$

thus, by Eqs. 4.51 and 4.49, we have the final results

$$G(x - x') = -(2\pi)^{-4} \int dk e^{-ik(x-x')}/k^2, \tag{4.54}$$

and

$$A^{\mu}(x) = \underset{0}{A}{}^{\mu}(x) - (2\pi)^{-4} \frac{1}{c} \int dk\, dx' j^{\mu}(x') e^{-ik(x-x')}/k^2. \tag{4.55}$$

Remark. We discuss now the above procedure in a more general form and context. We consider x and x' as continuously infinite matrix indices and a linear operator $\mathscr{L}$. In Eq. 4.48 $\mathscr{L}$ is the d'Alembertian $\Box^2$. The underlying vector space is an infinite-dimensional Hilbert space whose vectors we write in the form of kets; e.g., $|x\rangle$. The basis vectors $|x\rangle$ are orthogonal, but not normalized. We have $\langle x|x'\rangle = \delta^4(x - x')$. Note that the x's are four-vectors, and, although we use quantum-mechanical notation, everything is classical. The Green's functions satisfy then, in matrix notation, the equation

$$\mathscr{L}G = I, \tag{4.56}$$

i.e., the Green's operator G is the inverse of the operator $\mathscr{L}$. The Green's functions themselves are the matrix elements of G:[1]

$$G(x - x') = \langle x|G|x'\rangle. \tag{4.57}$$

Any inhomogeneous equation $\mathscr{L} f = \phi$, where f and ϕ are vectors in Hilbert space, with $\langle x|f\rangle = f(x)$, $\langle x|\phi\rangle = \phi(x)$, has the formal solution

$$f = f_0 + G\phi,$$

where f_0 is any null vector of the operator $\mathscr{L}$; i.e., $\mathscr{L} f_0 = 0$. This is precisely Eq. 4.49.

Assume now that we know a complete set of orthogonal eigenvectors of the operator G

$$G|k\rangle = g_k|k\rangle, \tag{4.58}$$

with

$$\langle x|k\rangle = \psi_k(x);$$

the index k is, in general, continuously infinite, in which case $\langle k|k'\rangle = \delta(k - k')$. We expand the vectors $|x\rangle$ in this new basis $|k\rangle$ and obtain

$$|x\rangle = \int |k\rangle\langle k|x\rangle\, dk.$$

---

[1] The quantum-mechanical bra, $\langle|$, and ket, $|\rangle$, and the matrix notation in Eq. 4.57 is just a convenient way of writing the eigenfunctions and their expansion; it has nothing to do with quantum mechanics.

If we use this expansion in Eq. 4.57, we have

$$G(x-x') = \int \langle x|G|k\rangle\langle k|x'\rangle\, dk = \int \langle x|k\rangle g_k \langle k|x'\rangle\, dk = \int dk\, g_k \psi_k(x)\psi^*{}_k(x'). \tag{4.59}$$

This is the important formula expressing Green's functions in terms of its complete set of orthogonal eigenfunctions. In our example, with $\mathscr{L} = \Box^2$,

$$\langle x|k\rangle = e^{-ikx},$$

$$g_k = -1/k^2,$$

and Eq. 4.59 reduces to Eq. 4.54.

Because Green's functions have a direct physical meaning as the potential induced on a point x due to a unit source on the point x', it is important to study their explicit forms and properties.

We first discuss the simple case of a static electric field which in itself will be a useful result later. Let the potentials A^μ be of the form

$$A^\mu : (\phi, 0, 0, 0),$$

and

$$\dot{\phi} = 0.$$

Equation 4.47 reduces in this case to the Poisson equation

$$\Delta\phi = -\frac{1}{c} j^0 = -\rho. \tag{4.60}$$

Using the techniques discussed before, we obtain

$$G(\mathbf{x}-\mathbf{x}') = (2\pi)^{-3} \int d\mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}/\mathbf{k}^2.$$

The integral on the right can be evaluated if we put $d\mathbf{k} = K^2 dK\, d\Omega$ $(\mathbf{x}-\mathbf{x}') = \mathbf{r}$, and $\mathbf{k}\cdot\mathbf{r} = KR\cos\theta$.

$$G(\mathbf{r}) = (2\pi)^{-3} \int dK e^{iKR\cos\theta} \sin\theta\, d\theta\, d\phi,$$

or[1]

$$G(R) = \frac{2}{R}(2\pi)^{-2} \int_0^\infty dK \sin(KR)/K = 1/4\pi R. \tag{4.61}$$

This is, of course, the well-known result for the static field at a distance R

---

[1] A useful formula for the conversion of a class of three-dimensional integrals into a one-dimensional one which is frequently met is

$$\int d^3k\, e^{i\mathbf{k}\cdot\mathbf{x}} f(K) = \frac{4\pi}{R} \int_0^\infty dK\, K f(K) \sin(KR). \tag{4.62}$$

Note also the integral $\int_0^\infty dK \sin(KR)/K = \pi/2$.

from a unit source. The Green's function depends only on the magnitude $R = |\mathbf{x} - \mathbf{x}'|$. The general solution of the Poisson Eq. 4.60 is then

$$\phi = \phi_0 - \int \frac{\rho(\mathbf{x}')\, d^3x'}{4\pi|x - x'|}, \tag{4.63}$$

where ϕ_0 is a solution of the Laplace equation, $\Delta\phi_0 = 0$, and will be determined by the boundary conditions. If we require, as usual, that $\phi \to 0$ as $\mathbf{x} \to \infty$, we find $\phi_0(x) \equiv 0$.

We now solve the Green's function, Eq. 4.54, in the general four-dimensional; i.e., nonstatic case. We write $dk = dk^0\, d^3k$. The integral in Eq. 4.54 can be evaluated either by carrying out the k^0 integration first or the $d^3\mathbf{k}$ integration first. In the former case,

$$G(x - x') = (2\pi)^{-4} \int d^3k \int_{-\infty}^{+\infty} dk^0\, e^{ik^0(x^0 - x^{0\prime})} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}/(k^{02} - \mathbf{k}^2).$$

We use the method of analytic continuation on the variable k^0. The integrand is a function of k^0 with two poles at

$$k^0 = \pm(\mathbf{k}^2)^{1/2} \equiv \pm K.$$

In the lower or upper half k^0 plane, depending on the sign of $(x^0 - x^{0\prime})$, the integrand goes to zero at infinity exponentially. Therefore, we can transform the k^0 integration into a contour integration. Various contours are possible giving rise to different possible solutions (Fig. 6) which, however, must differ from each other by the solutions of the homogeneous equations. Hence the different contours correspond to different boundary conditions.

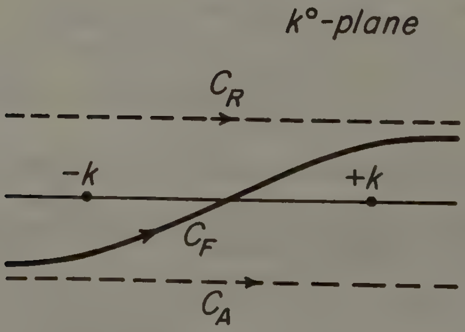


Fig. 6. Contours for k⁰—integration for various Green's functions

Retarded Green's Functions. Let us choose the contour C_R. We must close the contour in the upper half plane for $(x^0 - x^{0\prime}) < 0$, and in the lower half plane for $(x^0 - x^{0\prime}) > 0$. We denote the corresponding Green's functions by $D^{\mathrm{ret}}(x - x')$. In the former case the contour does not include any poles and the integral is zero; in the latter case the k^0 integration gives just the residue of the integrand at the two poles.

$$\int dk^0 e^{-ik^0(x^0 - x^{0\prime})}/(k^{02} - \mathbf{k}^2) = 2\pi i \frac{1}{2K}[e^{iK(x^0 - x^{0\prime})} + \\ + e^{-iK(x^0 - x^{0\prime})}]\theta(x^0 - x^{0\prime}),$$

where $\Theta\ (t)$ is the step function:

$$\begin{aligned} \theta(t) &= 0 \quad \text{for} \quad t < 0, \\ &= 1 \quad \text{for} \quad t > 0. \end{aligned} \tag{4.64}$$

Thus, with $|\mathbf{x} - \mathbf{x}'| = R$,

$$\begin{aligned} D^{\mathrm{ret}}(x - x') &= (2\pi)^{-4} 2\pi i \int dK K^2 e^{iKR\cos\theta} 2\pi \sin\theta\, d\theta [e^{iK(x^0 - x^{0\prime})} \\ &\quad + e^{-iK(x^0 - x^{0\prime})}]\theta(x^0 - x^{0\prime})/2K \\ &= (2\pi)^{-3}\pi i \int dK K [e^{iK(x^0 - x^{0\prime})} \\ &\quad + e^{-iK(x^0 - x^{0\prime})}]\theta(x^0 - x^{0\prime}) \int e^{iKRu} du \\ &= \frac{1}{4\pi R}\theta(x^0 - x^{0\prime})[\delta(x^0 - x^{0\prime} + R) + \delta(x^0 - x^{0\prime} - R)]. \end{aligned} \tag{4.65}$$

Because $R > 0$, only the second δ function contributes, in conjunction with the θ function, and we have finally

$$D^{\mathrm{ret}}(x - x') = \frac{1}{4\pi |\mathbf{x} - \mathbf{x}'|}\delta(x^0 - x^{0\prime} - |\mathbf{x} - \mathbf{x}'|). \tag{4.66}$$

The form of $D^{\mathrm{ret}}(x - x')$ clarifies the name "retarded functions" given to these solutions. The field at the "point" x due to the unit source at the point x' is different from zero only if the time of $\mathbf{x}'$ is earlier than the time at which one evaluates the field at $\mathbf{x}$. Moreover, we see from Eq. 4.66 that only signals travelling with a velocity of light from $\mathbf{x}'$ to $\mathbf{x}$ contribute to the field. To see this and the invariance property of the Green's function better, we rewrite Eq. 4.65 in the form

$$D^{\mathrm{ret}}(x - x') = (1/2\pi)\delta[(x - x')^2]\theta(x^0 - x^{0\prime}). \tag{4.67}$$

The function $\theta(x^0 - x^{0\prime})$ is invariant under inhomogeneous orthochronous Lorentz transformations, and so is $D^{\mathrm{ret}}(x - x')$. The invariant δ function $\delta[(x - x')^2]$ shows that D^{ret} is zero unless the vector $(x - x')$ is light-like. If we draw a light cone with x at its center, x' must lie *on* the backward light cone. (Fig. 7.) Later we shall make use of Eq. 4.67. In its derivation we have used the formula

$$\delta(t^2 - a^2) = \frac{1}{2a}[\delta(t-a) + \delta(t+a)], \tag{4.68}$$

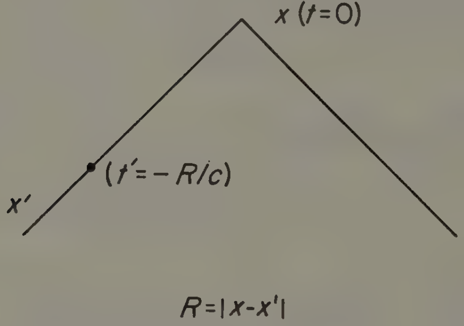


Fig. 7. Field point x ($t = 0$) **and the source point** x' **at retarded time** $t' = -R/c$.

which we prove here. The right-hand side of this equation is

$$(1/2a)\int dt f(t)[\delta(t-a) + \delta(t+a)] = (1/2a)[f(a) + f(-a)],$$

and the left-hand side can be transformed as follows:

$$\int_{-\infty}^{+\infty} dt\delta[(t-a)(t+a)]f(t) = \int_{-a-\varepsilon}^{-a+\varepsilon} dt\delta[(t-a)(t+a)]f(t) + \int_{a-\varepsilon}^{a+\varepsilon} dt\delta[(t-a)(t+a)]f(t).$$

If we set in the first term on the right $t + a = y$, and in the second $t - a = y$, we get

$$\int dy\delta[(y-2a)y]f(y) + \int dy\delta[(y+2a)y]f(y) = \frac{1}{2a}[\delta(t-a) + \delta(t+a)].$$

Q.E.D.

Equation 4.68 is a special case of the more general relation

$$\delta[g(a)] = \sum_k \frac{\delta(a - a_k)}{|g'(a_k)|}, \tag{4.69}$$

where a_k are the roots of the equation $g(a) = 0$. Other useful formulae for δ functions that we need are

$$\delta(-a) = \delta(a); \qquad \delta'(a) = -\delta(a)/a, \qquad \int_{-\infty}^{b} da\delta(a) = \theta(b).$$

Advanced Green's Functions. We now choose the contour C_A in Fig. 6, and call the corresponding solution the advanced Green's function, for in this case

$$D^{\text{adv}}(x-x')\begin{cases} =0 & \text{for} \quad x^0 > x^{0'} \\ \neq 0 & \text{for} \quad x^0 < x^{0'}. \end{cases}$$

Thus, $D^{\text{adv}}(x-x')$ represents a field at the point $\mathbf{x}$ at time t due to unit charge at $\mathbf{x}'$ at a *later* time t'. This type of solution may seem contrary to our physical idea of propagation with increasing time; however, due to the causal character of our equations, they are perfectly good solutions. This point will be made clearer where we use D^{adv}, later, to evaluate the field.

The evaluation of $D^{\text{adv}}(x-x')$ is exactly the same as before, and we find

$$D^{\text{adv}}(x-x') = (1/2\pi)\delta[(x-x')^2]\theta(x^{0'}-x^0),$$

or **(4.70)**

$$D^{\text{adv}}(x-x') = (1/4\pi R)\delta(x^0 - x^{0'} + R).$$

Again, D^{adv} fails to vanish only if $(x-x')$ is light-like *and* x' is on the forward light cone with respect to x.

The Green's functions D^{ret} and D^{adv} are highly singular functions, zero everywhere except on one side of the light cone, where they have a δ-function singularity. From their definitions we find that

$$D^{\text{ret}}(x-x') \to 0 \quad \text{as} \quad (x^0 - x^{0'}) \to +\infty,$$

$$\partial D^{\text{ret}}/\partial(x^0 - x^{0'}) \to 0 \quad \text{as} \quad (x^0 - x^{0'}) \to +\infty; \tag{4.71}$$

and

$$D^{\text{adv}}(x-x') \to 0 \quad \text{as} \quad (x^0 - x^{0'}) \to -\infty,$$

$$\partial D^{\text{adv}}/\partial(x^0 - x^{0'}) \to 0 \quad \text{as} \quad (x^0 - x^{0'}) \to -\infty.$$

The difference of the two Green's functions

$$D(x-x') = D^{\text{ret}}(x-x') - D^{\text{adv}}(x-x') \tag{4.72}$$

is a solution of the homogeneous equation. This function plays an important rôle in applications. We get

$$D(x-x') = (1/2\pi)\delta((x-x')^2)\varepsilon(x^0 - x^{0'}), \tag{4.73}$$

where $\varepsilon(t)$ is the so-called sign function

$$\varepsilon(t) = \begin{cases} +1 & \text{if} \quad t > 0, \\ -1 & \text{if} \quad t < 0, \end{cases} \tag{4.74}$$

and is invariant under orthochronous Lorentz transformations. Because

$$D^{\text{ret}}(-x) = D^{\text{adv}}(x), \tag{4.75}$$

the homogeneous Green's function $D(x)$ is an odd function

$$D(x'-x) = -D(x-x'). \tag{4.76}$$

Going back to the integral form of D^{ret} and D^{adv}, we derive a useful spectral representation for the D function: From

$$D = D^{\text{ret}} - D^{\text{adv}} = i(2\pi)^{-3}\int d^3k e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}[e^{iK(x^0-x^{0\prime})} + e^{-iK(x^0-x^{0\prime})}] \\ \times \varepsilon(x^0-x^{0\prime})/2K,$$

and, using the identity

$$e^{\pm ik(x^0-x^{0\prime})} = \int dk^0\delta(k^0 \pm K)e^{-ik^0(x^0-x^{0\prime})},$$

we find

$$D(x-x') = i(2\pi)^{-3}\int dk e^{-ik(x-x')}[\delta(k^0+K)+\delta(k^0-K)]\varepsilon(x^0-x^{0\prime})/2K;$$

or, with Eq. 4.67,

$$D(x-x') = i(2\pi)^{-3}\int dk e^{-ik(x-x')}\delta(k^2)\varepsilon(k^0). \tag{4.77}$$

This equation will allow us to show that the time derivative of the D function at the origin is a three-dimensional δ function.

$$\left.\frac{\partial D(x-x')}{\partial x^0}\right|_{x^0=x^{0\prime}} = (2\pi)^{-3}\int d\mathbf{k}e^{i\mathbf{k}(\mathbf{x}-\mathbf{x}')}\int_{-\infty}^{+\infty} dk^0k^0\delta(k^2)\varepsilon(k^0).$$

The integral over k^0 gives just unity, because

$$\int dk^0|k^0|[\delta(k^0-K)+\delta(k^0+K)]/2K = 1;$$

hence

$$\partial D(x-x')/\partial(x^0-x^{0\prime})|_{x^0=x^{0\prime}} = \delta(\mathbf{x}-\mathbf{x}'). \tag{4.78}$$

This equation can be written covariantly in the form

$$\left.\int_\sigma d\sigma_\mu \frac{\partial D(x-x')}{\partial x_\mu} f(x)\right|_{x'\varepsilon\sigma} = f(x'). \tag{4.78$'$}$$

Sometimes it may be convenient to introduce a covariant "spatial" δ function δ^μ

$$\partial D/\partial x_\mu = \delta^\mu(x-x') = n^\mu\delta(\mathbf{x}-\mathbf{x}'),$$

such that

$$\int_\sigma d\sigma_\mu\delta^\mu(x-x')f(x) = f(x')$$

when both x and x' are on the space-like surface σ.

Other Invariant Green's Functions. Referring back to Eq. 4.54 and Fig. 6, there are two other important Green's functions. One is defined by the contour C_F. Depending on the sign of $(x^0 - x^{0\prime})$, we get the contribution of one or the other pole only.

$$D_F(x - x') = \theta(x^0 - x^{0\prime})D^+(x - x') - \theta(x^{0\prime} - x^0)D^-(x - x'), \quad \textbf{(4.79)}$$

where D^+ and D^- are the contributions of the poles at $+K$ and $-K$, respectively, and are given by

$$D^+(x) = i(2\pi)^{-3} \int d^3\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}e^{-iKx^0}/2K = i(2\pi)^{-3} \int dk e^{-ikx}\delta(k^2)\theta(k^0), \quad \textbf{(4.80)}$$

and

$$D^-(x) = i(2\pi)^{-3} \int d^3k e^{i\mathbf{k}\cdot\mathbf{x}}e^{iKx^0}/2K = -i(2\pi)^{-3} \int dk e^{-ikx}\delta(k^2)\theta(-k^0). \quad \textbf{(4.81)}$$

It follows from Eq. 4.77 that

$$D = D^{\text{ret}} - D^{\text{adv}} = D^+ + D^-. \quad \textbf{(4.82)}$$

We also have the properties

$$D^+(-x) = -D(x),$$

and

$$(D^+)^* = D^-, \quad \textbf{(4.83)}$$

where the * denotes complex conjugation.

The other Green's function we want to introduce is obtained by taking the principal value of the k^0 integration. We shall denote this function by $\bar{D}$. Because the principal value of a singular integral is equal to 1/2 the residue of the pole term, we obtain immediately

$$\bar{D}(x) = \tfrac{1}{2}(D^{\text{ret}} + D^{\text{adv}}). \quad \textbf{(4.84)}$$

This function is even and real:

$$\bar{D}(x) = \bar{D}(-x) = [\bar{D}(x)]^*. \quad \textbf{(4.85)}$$

We have introduced four invariant Green's functions D^{ret}, D^{adv}, D_F and $\bar{D}$. All these solutions differ by a solution of the homogeneous equation. Thus, in addition to Eq. 4.72, we have the relations

$$\begin{aligned} D^{\text{ret}} - \bar{D} &= D/2, \\ D^{\text{adv}} - \bar{D} &= -D/2, \quad \textbf{(4.86)} \\ D_F - \bar{D} &= \tfrac{1}{2}(D^+ - D^-) \equiv (1/2)D^1(x). \end{aligned}$$

In the last equation we have introduced a second (besides D) linearly independent solution of the homogeneous equation, $D^1(x)$.

Green's Functions for the Scalar Field. The previous theory of Green's functions can be applied immediately to the case of a single scalar field $\phi(x)$, satisfying an equation of the form

$$(\Box^2 + m^2)\phi(x) = f(x), \tag{4.87}$$

where $f(x)$ is the source of the scalar field. In effect, this case is simpler than that of electrodynamics, for we do not have any subsidiary condition as was the case before [see Eqs. 4.46 and 4.47].

From the general expression for the Green's function Eq. 4.59, with $G = (\Box^2 + m^2)^{-1}$, we obtain immediately

$$G(x - x') = (2\pi)^{-4} \int dk e^{-ik(x-x')}/(-k^2 + m^2). \tag{4.88}$$

The discussion of this equation is now very similar to that of Eq. 4.54. The poles in the k^0 plane are now at

$$k^0 = \pm(\mathbf{k}^2 + m^2)^{1/2}.$$

This is the only difference. Exactly the same type of Green's function with the same contours can be defined here and will be denoted by

$$\Delta^{\text{ret}}, \Delta^{\text{adv}}, \bar{\Delta}, \Delta_F,$$

and the corresponding solutions of the homogeneous equation by

$$\Delta^+, \Delta^-, \Delta, \Delta^1.$$

All the relations between these functions are exactly the same as between the corresponding D functions. In the spectral representations, $\delta(k^2)$ must be replaced by $\delta(k^2 - m^2)$.

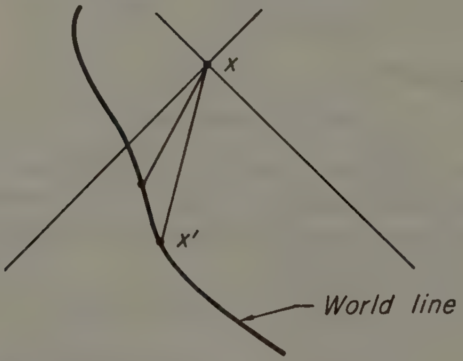


Fig. 8. For massive particles emitted from the source the entire world line of the source particle inside the backward light cone contributes to the field at x.

As a result of the nonvanishing mass, the invariant Δ functions are also different from zero outside the light cone, as we shall see. For example, Δ^{ret} is different from zero inside the backward line cone and Δ^{adv} inside the forward light cone. This implies that the field at the point x is not only determined by the intersection of the world line of the particle with the backward line cone, but by the whole history of the world line inside the backward line cone. Of course, this means that the velocity of the signal from the source to the field point is less or equal to the velocity of light. Because the signal is carried now by particles with nonzero rest mass, this is what one would expect to happen. (Fig. 8.)

To evaluate these singular functions explicitly, we notice that the defining integral [see Eq. 4.80]

$$\Delta^+(x) = i(2\pi)^{-3} \int dk e^{-ikx} \delta(k^2 - m^2)\theta(k^0)$$

converges if we replace x by $x - i\xi$, where ξ is a four-vector in the upper light cone. One can then define $\Delta^+(x)$ as the limit, when $\xi \to 0$, of a function which is well defined. The integral

$$i(2\pi)^{-3} \int dk e^{-ik(x-i\xi)} \delta(k^2 - m^2)\theta(k^0)$$

being an invariant function, is a function of $(x - i\xi)^2$ and m only. Hence we can evaluate it, for example, by taking $x = (x^0, 0, 0, 0)$ and $\xi = (\xi^0, 0, 0, 0)$.

$$\begin{aligned}
\Delta^+(x) &= \lim_{\xi^0 \to 0} i(2\pi)^{-3} \int dk e^{-ik^0(x^0 - i\xi^0)} \delta(k^2 - m^2)\theta(k^0) \\
&= \lim_{\xi^0 \to 0} i(2\pi)^{-3} \int dk^0 e^{-ik^0(x^0 - i\xi^0)} \theta(k^0) \int 2\pi \frac{k^2\, dk \delta[(k_0^2 - m^2)^{1/2} - k]}{k} \\
&= \lim_{\xi^0 \to 0} i(2\pi)^{-2} \int_m^\infty dk^0 (k_0^2 - m^2)^{1/2} e^{-ik^0(x^0 - i\xi^0)} \\
&= i \frac{m^2}{8\pi} \lim_{\xi^0 \to 0} \frac{H_1^{(2)}(m(x^0 - i\xi^0))}{m(x^0 - i\xi^0)},
\end{aligned}$$

where $H_1^{(1)}$ is the Hankel function of the first kind.[1] In an arbitrary frame, $(x^0 - i\xi^0)$ must be replaced by $[(x - i\xi)^2]^{1/2}$.

---

[1] We use the general formula

$$H_\nu^{(\frac{1}{2})} = \mp \frac{2i(z/2)^{-\nu}}{\sqrt{\pi}\Gamma(\frac{1}{2} - \nu)} \int_1^\infty e^{\pm izt}(t^2 - 1)^{-\nu-\frac{1}{2}} dt, \quad Im\, z > 0, \quad Re\, \nu < 1/2,$$

(i.e., upper sign for $H_\nu^{(1)}$, lower sign or $H_\nu^{(2)}$). [H. Bateman, *Higher transcendental function* (McGraw-Hill Book Company, Inc., 1953) Vol. II, p. 83.]

Using Eqs. 4.82 and 4.83, we obtain

$$\Delta(x) = \Delta^{+} + \Delta^{-} = 2\,Re\,\Delta^{+}(x) = \varepsilon(x)\left[\frac{\delta(x^2)}{2\pi} - \theta(x^2)\,\frac{m^2}{4\pi}\,\frac{J_1[m(x^2)^{1/2}]}{m(x^2)^{1/2}}\right],$$

and

$$\Delta^{1}(x) = 2\,Im\,\Delta^{+}(x) = \frac{im^2}{4\pi}\,Im\,\frac{H_1^{(2)}[m(x^2)^{1/2}]}{m(x^2)^{1/2}}\,.$$

The second term in $\Delta(x)$ and $\Delta'(x)$ can be expanded if $m(x^2)^{1/2}$ is small, and we find

$$\Delta(x) = \frac{\delta(x^2)}{2\pi}\,\varepsilon(x) + \frac{\theta(x^2)}{2\pi}\left[\frac{m^2}{4}\,\varepsilon(x) + \frac{x^2m^4}{32}\,\varepsilon(x) + \cdots\right],$$

and

$$\Delta'(x) = -\frac{2i}{4\pi^2x^2} + \frac{i}{4\pi^2}\left[\,m^2\ell n\left(\frac{\gamma}{2}\right)m(x^2)^{1/2} - \frac{m^2}{2} + \cdots\right].$$

Comparing this with Eq. 4.73 we see that the first term of $\Delta(x)$ is equal to $D(x)$, the second term is nonzero for $x^2 > 0$ and shows the behavior of the function inside the light cone.

There are invariant Green's functions for other field equations; they can be discussed, in principle, along the same lines.

Cauchy Problem. For the general solution of the inhomogeneous field equation, we need to know the solutions of the homogeneous equation; see, for example, Eq. 4.49. We have indicated that the solutions of the homogeneous equation will be determined by suitable boundary conditions. In connection with this, we give now a general procedure for writing down these solutions.

The problem is to find the solution of a linear, second-order, partial-differential equation everywhere, if the function and its first derivative is given on a space-like surface.

For the linear hyperbolic differential equation,

$$\Box^2 A^{\mu} = 0,$$

the solution of the Cauchy problem can be written in terms of the D functions discussed above:

$$A^{\mu}(x') = \int_{\sigma} d\sigma_{\nu}[\bar{A}^{\mu}(x)\,\partial^{\nu}D(x-x') - D(x-x')\bar{A}^{\mu,\nu}(x)], \qquad \textbf{(4.89)}$$

where $\bar{A}^{\mu}$ and $\bar{A}^{\mu,\nu}$ are the given functions on the surface σ. Indeed $A^{\mu}(x)$ given by this equation satisfies $\Box^2 A^{\mu} = 0$, because D itself is a solution of this equation. We must show now that on the prescribed space-like surface σ, $A^\mu(x)$ and its derivatives take the prescribed values. Noting that $D(x - x') = 0$, if x' is on the same surface σ as x, we get

$$A^\mu(x')|_{x'\epsilon\sigma} = \int_\sigma d\sigma_\nu \bar{A}^\mu(x)\, \partial^\nu D(x - x')|_{x'\epsilon\sigma},$$

which, by Eq. 4.78′, is precisely $\bar{A}^\mu(x')$. Similarly,

$$A^{\mu,\rho}(x')|_{x'\epsilon\sigma} = -\int_\sigma d\sigma_\nu[\bar{A}^\mu(x)\, \partial^{\nu\rho}D(x - x') - \partial^\rho D(x - x')\bar{A}^{\mu,\nu}(x)]|_{x'\epsilon\sigma}$$

which reduces to $A^{\mu,\rho}(x')$.

In a similar way, the solution of the Cauchy problem for the scalar field,

$$(\Box^2 - m^2)\phi(x) = 0,$$

is given by

$$\phi(x') = \int_\sigma d\sigma_\mu[\Delta(x - x')\bar{\phi}^{,\mu}(x) - \bar{\phi}(x)\, \partial^\mu\Delta(x - x')],$$

where $\bar{\phi}(x)$ and $\bar{\phi}^{,\mu}(x)$ are the given values of the function and its derivative on the space-like surface σ.

The solution to the Cauchy problem is unique, for if we had two distinct solutions, their difference would be one also. But this difference and its derivative vanishes everywhere on a space-like surface. Then the second-order differential equation $\Box^2 A^\mu = 0$ or $(\Box^2 - m^2)\phi = 0$ shows that it must vanish everywhere in space-time.

Solutions of the Inhomogeneous Equation. Having discussed the Green's functions and the solutions of the homogeneous equation for a prescribed data at a given time (space-like surface σ), we can write now the general solution of the inhomogeneous equation, which is really our problem.

For electrodynamics we go back to Eq. 4.49. Let us choose the Cauchy data to be on the surfaces $t = -\infty$ or $t = +\infty$. Let the corresponding solutions of the homogeneous equation be denoted by $A_{\text{in}}(x)$ and $A_{\text{out}}(x)$, respectively. For these two cases, we must use the retarded and the advanced Green's functions, and we get

$$A^\mu(x) = A^\mu_{\text{in}}(x) + \frac{1}{4\pi c}\int \frac{\delta(x^0 - x^{0'} - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|}\, j^\mu(x')\, dx',$$

$$A^\mu(x) = A^\mu_{\text{out}}(x) + \frac{1}{4\pi c}\int \frac{\delta(x^0 - x^{0'} + |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|}\, j^\mu(x')\, dx'. \qquad \mathbf{(4.90)}$$

In the first case, we determine the field at the point x from the past, starting with the initial value at $t = -\infty$ and in the second case, from the future backwards starting with its initial value at $t = +\infty$. Because the development of the system is causal, these two methods are equivalent.

The two solutions $A^{\mu}_{\text{in}}(x)$ and $A^{\mu}_{\text{out}}(x)$ are related by a third solution of the homogeneous equation; from Eq. 4.72, we get

$$A^{\mu}_{\text{out}}(x) = A^{\mu}_{\text{in}}(x) + \frac{1}{c}\int D(x - x')j^{\mu}(x')\,dx'. \tag{4.91}$$

If measurements are made at distant past and distant future, where the field is undisturbed, and then a disturbance is introduced at finite times, as in *scattering experiments*, the integral in Eq. 4.91 determines the net effect of the interaction. In quantum phenomena, this is the only accessible maseurement, and the "ratio" of A_{out} to A_{in} is the so-called "*scattering matrix*":

$$A^{\mu}_{\text{out}}(x) = SA^{\mu}_{\text{in}}(x). \tag{4.92}$$

Thus, even classically, it is meaningful to introduce the scattering "matrix" which is well defined by Eqs. 4.92 and 4.91.

Subsidiary Condition. For the scalar field interacting with a current, the solution given in the previous paragraph completes the problem. However, in the case of electrodynamics, we have worked in a particular gauge such that $A^{\mu}_{,\mu} = 0$.

$$A^{\mu}_{,\mu} = 0.$$

It will be shown now that if this subsidiary condition is satisfied for the initial free fields $A^{\mu}_{\text{in}}(x)$ or $A^{\mu}_{\text{out}}(x)$ in Eq. 4.90, then it is satisfied always.

Let us assume now that

$$A^{\mu}_{\text{in},\mu} = 0, \qquad A^{\mu}_{\text{out},\mu} = 0;$$

then, by Eq. 4.90

$$A^{\mu}_{,\mu} = \frac{1}{c}\int \frac{\partial D^{\substack{\text{adv}\\ \text{ret}}}(x - x')}{\partial x^{\mu}}\, j^{\mu}(x')\,dx' = \frac{1}{c}\int \frac{\partial D^{\substack{\text{adv}\\ \text{ret}}}(x - x')}{\partial x'^{\mu}}\, j^{\mu}(x')\,dx',$$

or, integrating by parts

$$A^{\mu}_{,\mu} = -\frac{1}{c}\int dx'\,\frac{\partial}{\partial x'^{\mu}}\left(D^{\substack{\text{adv}\\ \text{ret}}}(x - x')j^{\mu}(x')\right) + \frac{1}{c}\int dx' D^{\substack{\text{adv}\\ \text{ret}}}(x - x')j^{\mu}_{,\mu}(x').$$

The second term vanishes by the equation of continuity $j^{\mu}_{,\mu} = 0$. The first term can be transformed into a surface integral, passing through the point x,

$$A^{\mu}_{,\mu} = -\frac{1}{c}\int_{\sigma}^{\sigma} d\sigma'_{\mu} D^{\overset{\text{adv}}{\text{ret}}}(x - x')j^{\mu}(x'),$$

which vanishes, because $D^{\overset{\text{adv}}{\text{ret}}}(x - x') = 0$ for space-like separations.

An alternate simple proof is the following: Let $A^{\mu}_{,\mu} \equiv x$, which satisfies

$$\Box^2 x = 0.$$

Now, if χ and $\chi_{,\mu}$ are zero on a space-like surface, then $\chi = 0$ everywhere.

5. *Further Developments and Exercises*

1. Using the invariance of the Lagrangian for the interaction of fields with external currents and particles, show that the total angular momentum is conserved.

2. Using the action integral, Eq. 4.21, for a charged particle interacting with an electromagnetic field, derive the coupled equations of motion and the conservation law for the energy-momentum tensor.

3. Consider the force between two uniformly moving charged particles (for example, one moving along the x axis, the other along the y axis) and show that the law "action is equal to reaction" does not hold and interpret this result using the conservation of total momentum of the particles plus the field.

4. Evaluate the energy-momentum tensor $T^{\mu\nu}$ of interacting scalar and electromagnetic fields given by the Lagrangian, Eq. 4.29, and interpret the conservation of total energy momentum in terms of individual energy momenta of the fields and the energy-momentum exchange between the fields.

5. Show that Eq. 4.36 is the most general isotropic tensor of rank 4.

6. Evaluate the current vector and the energy-momentum tensor for vector fields, and discuss their interpretation.

7. Show by the method of Chapter I, Sect. 5, that the 2 × 2 unitary unimodular matrices u form a two-dimensional representation of the three-dimensional rotation group. Therefore, a field that is a two-component spinor in the internal space transforms as

$$\psi'_{\alpha} = u_{\alpha}{}^{\beta}\psi_{\beta}.$$

8. For the isospinor field of Problem 7 and the isovector field transforming as

$$\psi'_{\alpha} = R_{\alpha}{}^{\beta}\psi_{\beta},$$

where R is the three-dimensional rotation matrix, calculate the generator of the infinitesimal transformation, assuming the Lagrangian is invariant under these transformations.

9. Discuss the introduction of the Yang-Mills and the gauge fields, starting from current conservation and gauge transformations of the second kind of the type of Eq. 4.50.

10. Prove Eq. 4.69 by first showing that

$$\delta(g'(a_k)(a - a_k)) = \delta(a - a_k)/g'(a_k).$$

11. Relate the principal-value integral to the residue of the pole, and prove Eq. 4.84.

12. Solve the equation $(\Delta^2 - \lambda)\phi(\mathbf{r}) = \rho(\mathbf{r})$ by the method of Green's functions. In particular, find the scalar mesonic field of a point nuclear charge.

13. Find the Green's functions for the Dirac equation.

BIBLIOGRAPHY FOR CHAPTER IV

In addition to books quoted at the end of Chapter III, the following articles should be consulted for some of the special topics discussed in this chapter.

1. S. Kahana and H. R. Coish, "Classical Meson Theory I and II," *Amer. Jour. of Phys. 24*, (1956) 225 and 390.
2. For a detailed study of classical aspects of various meson fields, see the series of papers by P. Havas, *Phys. Rev.*, *87*, (1952), 309; *91*, (1953), 997; *91*, (1953), 393; *93*, (1954), 882; *93*, (1954), 1400; *113*, (1959), 732; *116*, (1959), 202.
3. H. C. Corben and J. Schwinger, "The Electromagnetic Properties of Mesotrons," *Phys. Rev.*, *58*, (1940), 953.
4. C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.*, *96*, (1954), 191.
5. J. Sakurai, "Theory of Strong Interactions," *Ann. of Physics*, *11*, (1960), 1.

V.

RADIATION AND RADIATION REACTION

In this Chapter we concentrate, in detail, on the electromagnetic radiation. First, the field of a moving charged particle is discussed; then we define the radiation field and investigate its properties. This is followed by a discussion of the effect of the radiation field on the charged particle and the modified equations of the motion of the particle with radiation reaction which present unusual features. Additional requirements must be used to obtain physically meaningful solutions. Finally we discuss the classical mass renormalization and the problem of the electromagnetic nature of the mass of the electron. In the application of the results to the real physical situation, particularly when one speaks of the nature of the electron, quantum effects must be taken into effect. But the problems discussed here are fundamental for the consistency of the classical theory and what we may expect when the theory is quantized.

1. *Radiation Field of a Moving Particle*

Lienard-Wiechert Potentials

Let us put the external free field $A^{\mu}_{\substack{\text{in}\\\text{out}}}(x) = 0$. We want to evaluate the field produced by the particle alone whose current density is given by

$$j^{\mu} = ec \int_{-\infty}^{+\infty} \dot{z}^{\mu}(s)\, \delta(x - z(s))\, ds$$

where $\dot{z}^\mu(s)$ is the four-velocity of the particle and s the proper time. The field can be evaluated from Eq. 4.90 and we get

$$A^\mu = e \iint D^{\mathrm{ret}}(x - x')\dot{z}^\mu(s)\, \delta(x' - z(s))\, ds\, dx'. \tag{5.1}$$

We use Eq. 4.67 and carry out the x'-integration

$$A^\mu(x) = \frac{e}{2\pi} \int_{-\infty}^{+\infty} ds\, \delta[(x - z(s))^2]\dot{z}^\mu(s)\theta(x^0 - z^0). \tag{5.1$'$}$$

The argument of the δ-function, $(x - z(s))^2$, has two roots with respect to s; the θ-function picks up the root with $x^0 > z^0(s)$. If we note that

$$\frac{d}{ds}[(x - z(s))^2] = -2(x - z(s))^\sigma \dot{z}_\sigma \tag{5.2}$$

and use Eq. 4.69, we can integrate Eq. 5.1′ and find

$$A^\mu(x) = \frac{e}{4\pi} \frac{\dot{z}^\mu}{(x - z)^\sigma \dot{z}_\sigma}\bigg|_{s=s_0} \tag{5.3}$$

These are the Lienard-Wiechert potentials. The right-hand side has to be evaluated at the time s_0 such that $x - z(s_0)$ is light-like and $x^0 - z^0(s_0) = R > 0$. This is just the point where the trajectory of the particle cuts the light cone with vertex at x. (Fig. 9.) Only from the point $z(s_0)$ can a

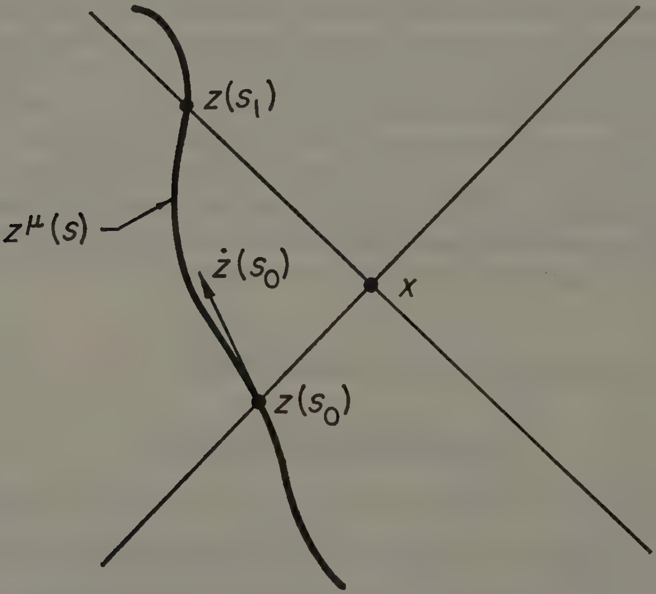


Fig. 9. World line $z^\mu(s)$, source points for retarded and advanced potentials and the velocity $\dot{z}^\mu(s_0)$.

light signal reach the point (x, t): The three-dimensional distance between particle and the field point is $|x - z(s_0)| = R$ and the time interval $t - t'(s_0) = R/c$.

Had we used an advanced function the only point of the trajectory of the particle which contributes to the field at x would be $z(s_1)$ (Fig. 9).

In the case of the scalar field, the Green's functions $\Delta^{\text{adv}}_{\text{ret}}(x - x')$ are also different from zero inside the light cone. Therefore, referring to Fig. (8) the whole trajectory inside the light-cone contributes to the field at x. This corresponds to signal propagating with a velocity less or equal the velocity of light. In quantum mechanical language, if light signals are interpreted as exchange of massless photons, in the case of scalar field we can talk of exchange of particles with mass, mesons.

We now write the Lienard-Wiechert potentials, Eq. 5.3, in a more familiar form. Introducing the velocity of the particle v, the distance

$$\mathbf{R} = \mathbf{x} - \mathbf{z}(s_0), \qquad \mathbf{n} = \mathbf{R}/R \tag{5.4}$$

we have

$$\dot{z}^\mu : (c, \mathbf{v}) \frac{dt}{ds}$$

$$(x - z)^\sigma \dot{z}_\sigma = (Rc - \mathbf{R} \cdot \mathbf{v}) \frac{dt}{ds};$$

and Eq. 5.3 can be split into a scalar and vector potentials

$$\phi = \frac{e}{4\pi R} \frac{1}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})} \tag{5.5}$$

$$\mathbf{A} = \frac{e}{4\pi R} \frac{\boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})} \tag{5.5$'$}$$

where

$$\boldsymbol{\beta} = \mathbf{v}/c.$$

Note that the velocity of the particle refers to the *retarded time* $t' = t - R/c$; and R is the distance between the field point $\mathbf{x}$ and the position of the particle at the retarded time t'.

As a special example, we consider a particle at rest.

$$\begin{aligned} &z^\mu : (s, 0, 0, 0) \\ &\dot{z}^\mu : (1, 0, 0, 0) \end{aligned} \tag{5.6}$$

and we obtain

$$\mathbf{A} = 0 \quad \text{and} \quad \phi = e/4\pi R.$$

This is the simple Coulomb field.

The Field Tensor $F^{\mu\nu}$

In evaluating $F^{\mu\nu}$ it is more convenient to go back to the integral expression, Eq. 5.1, rather then differentiate the result, Eq. 5.3. We have

$$A^{\mu,\nu}(x) = e\int ds \frac{\partial}{\partial x_\nu}[D^{\mathrm{ret}}(x-z)]\dot{z}^\mu.$$

Since $D^{\mathrm{ret}}(x)$ is a function of x^2, we get

$$\begin{aligned} A^{\mu,\nu} &= e\int ds \frac{\partial D^{\mathrm{ret}}}{\partial[(x-z)^2]}\frac{\partial[(x-z)^2]}{\partial x_\nu}\dot{z}^\mu \\ &= 2e\int ds (x-z)^\nu \dot{z}^\mu \frac{dD^{\mathrm{ret}}}{ds}\frac{ds}{d[(x-z)^2]} \end{aligned}$$

or, using Eq. 5.2,

$$A^{\mu,\nu} = -e\int ds \frac{(x-z)^\nu \dot{z}^\mu}{(x-z)^\sigma \dot{z}_\sigma}\frac{dD^{\mathrm{ret}}}{ds}.$$

We integrate this equation by parts

$$A^{\mu,\nu} = -e\left[D^{\mathrm{ret}}\frac{(x-z)^\nu \dot{z}^\mu}{(x-z)^\sigma \dot{z}_\sigma}\right]_{-\infty}^{+\infty} + e\int ds D^{\mathrm{ret}}(x-z)\frac{d}{ds}\left[\frac{(x-z)^\nu \dot{z}^\mu}{(x-z)^\sigma \dot{z}_\sigma}\right].$$

The first term is zero. In the second term we replace D^{ret} by $\delta[(x-z)^2]$ and use again Eq. 5.2 and integrate

$$A^{\mu,\nu} = \frac{e}{4\pi}\frac{1}{(x-z)^\sigma \dot{z}_\sigma}\frac{d}{ds}\left[\frac{(x-z)^\nu \dot{z}^\mu}{(x-z)^\sigma \dot{z}_\sigma}\right].$$

With

$$\begin{aligned} (x-z)^\sigma \dot{z}_\sigma &= R \\ (x-z)^\sigma \ddot{z}_\sigma &= Q \end{aligned} \tag{5.7}$$

the final result is

$$\begin{aligned} F^{\mu\nu} = \frac{e}{4\pi}\frac{1}{R^3}[&(x-z)^\mu \ddot{z}^\nu R - (x-z)^\mu \dot{z}^\nu Q + (x-z)^\mu \dot{z}^\nu \\ &- (x-z)^\nu \ddot{z}^\mu R + (x-z)^\nu \dot{z}^\mu Q - (x-z)^\nu \dot{z}^\mu]. \end{aligned} \tag{5.8}$$

Before we discuss this equation in detail let us look at two simple examples. For the Coulomb-field with Eq. 5.6 and $R = x^0 - z^0$ and $(x-z)^\mu = (x^\mu - s\delta^{\mu 0})$ we obtain

$$F^{\mu\nu} = \frac{e}{4\pi R^3}[(x^\mu - s\,\delta^{\mu 0})\,\delta^{\nu 0} - (x^\nu - s\,\delta^{\nu 0})\,\delta^{\mu 0}]$$

or,

$$\mathbf{B} = 0, \qquad \mathbf{E} = -\frac{e}{4\pi R^3}\mathbf{R}. \tag{5.9}$$

As a second example, consider a uniformly moving particle

$$\ddot{z} = 0; \qquad \dot{z} = (c, v, 0, 0)\frac{dt}{ds} = \gamma(1, \beta, 0, 0)$$

then

$$R = (x - z)^\sigma \dot{z}_\sigma = (x^0 - z^0)c\frac{dt}{ds} + (x^1 - z^1)v\frac{dt}{ds}$$

and we get

$$F^{\mu\nu} = \frac{e}{4\pi R^3}[(x^0 - z^0)\,\delta^{\mu 0}\,\delta^{\nu 1}\beta\gamma + (x^0 - z^0)\,\delta^{\mu 1}\,\delta^{\nu 0} - (\mu - \nu)]$$

$$= \frac{e\gamma}{4\pi R^2}(1 - \beta)(\delta^{\mu 1}\,\delta^{\nu 0} - \delta^{\nu 1}\,\delta^{\mu 0}) \tag{5.9$'$}$$

or

$$E_x = \frac{e\gamma}{4\pi R^2}(1 - \beta); \qquad E_y = E_z = 0\,.$$

The field $F^{\mu\nu}$ given by Eq. 5.8 consists of two parts:

(a) Field depending on the velocities $\dot{z}$ alone and proportional to $1/R^2$—*near field*

(b) Field depending on the acceleration $\ddot{z}$ and proportional to $1/R$—*far field* or *radiation.*

Let us look at these two terms in terms of the electric and magnetic fields.

$1/R^2$-field:

$$\mathbf{E}(\dot{z}) = \frac{e}{4\pi R^2}\frac{(1 - \beta^2)(\mathbf{n} - \boldsymbol{\beta})}{(1 - \mathbf{n}\cdot\boldsymbol{\beta})^3}$$

$$\mathbf{B}(\dot{z}) = \frac{e}{4\pi R^2}\frac{(1 - \beta^2)\boldsymbol{\beta} \times \mathbf{n}}{(1 - \mathbf{n}\cdot\boldsymbol{\beta})^3} \tag{5.10}$$

$1/R$-field:

$$\mathbf{E}(\ddot{z}) = \frac{e}{4\pi cR}\frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \mathbf{n}\cdot\boldsymbol{\beta})^3},$$

$$\mathbf{B}(\ddot{z}) = \frac{e}{4\pi cR}\frac{(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}} + \mathbf{n}(\mathbf{n}\cdot\boldsymbol{\beta} \times \dot{\boldsymbol{\beta}})}{(1 - \mathbf{n}\cdot\boldsymbol{\beta})^3}, \tag{5.11}$$

where $\boldsymbol{\beta} = (1/c)\mathbf{v}$ and the velocity $\mathbf{v}$ refers to the retarded time.

The radiation field, Eq. 5.11, simplifies for a slowly moving particle, $\beta \ll 1$, to

$$\begin{aligned} \mathbf{E}(\ddot{z}) &\cong \frac{e}{4\pi cR}\,\mathbf{n}\times(\mathbf{n}\times\dot{\boldsymbol{\beta}}) \\ \mathbf{B}(\ddot{z}) &\cong \frac{-e}{4\pi cR}\,\mathbf{n}\times\dot{\boldsymbol{\beta}} \\ &= \mathbf{n}\times\mathbf{E}(\ddot{z}) \end{aligned} \tag{5.12}$$

showing that electric and magnetic fields and the direction $\mathbf{n}$ which connects the field point to the position of the particle at retarded time are mutually orthogonal.

The Poynting vector of the radiation field is given by

$$\mathbf{S} = c\mathbf{E}(\ddot{z})\times\mathbf{B}(\ddot{z}) = \frac{e^2}{(4\pi)^2 cR^2}\,[\mathbf{n}\times(\mathbf{n}\times\dot{\boldsymbol{\beta}})]\times[\mathbf{n}\times\dot{\boldsymbol{\beta}}].$$

The energy flux is radial and is in the direction $\mathbf{R}$. The value of the energy flux in the solid angle $d\Omega$ is independent of R and equals

$$\frac{e^2}{(4\pi)^2 c}\,d\Omega(\mathbf{n}\times\dot{\boldsymbol{\beta}})^2.$$

To find the energy radiated per unit time we integrate over the sphere of radius R. With $\mathbf{n}\times\boldsymbol{\beta} = (1/c)\dot{\mathbf{v}}\sin\theta$ we have

$$\begin{aligned} \int \mathbf{S}\cdot\mathbf{df} &= \frac{e^2}{(4\pi)^2 c^3}\,\dot{v}^2 2\pi \int \sin^3\theta\,d\theta \\ &= \frac{2}{3}\frac{e^2}{4\pi}\frac{\dot{v}^2}{c^3} \end{aligned} \tag{5.13}$$

In the following sections these and other properties of the radiation field will be investigated in detail and quite generally without making the simplifying assumption, $\beta \ll 1$, used above.

2. *Properties of the Radiation Field*

Null Fields

We start the discussion with the field $F^{\mu\nu}$ produced by a moving particle given by Eq. 5.8. This field satisfies the free Maxwell equations

$$\partial F^{\mu\nu}/\partial x^\nu = 0$$

everywhere except on the world line of the particle. Consequently we have the conservation law for the electromagnetic energy momentum tensor

$$\partial T^{\mu\nu}/\partial x^\nu = 0$$

everywhere except on the world line.

It follows from Eq. 5.8 that the field of the moving particle outside the world line satisfies the properties

$$F_{\mu\nu}\varepsilon^{\mu\nu\sigma\rho}F_{\sigma\rho} = 0 \tag{5.14}$$

and

$$F_{\mu\nu}\varepsilon^{\rho\sigma\mu\nu}(x - z)_\sigma = 0. \tag{5.15}$$

These relations follow from the fact that $\varepsilon^{\mu\nu\sigma\rho}$ is antisymmetric whereas the products $F_{\mu\nu}F_{\sigma\rho}$ and $F_{\mu\nu}(x - z)_\sigma$ are symmetric in the indices $(\mu\sigma)$ and $(\nu\sigma)$, respectively, as can be seen by Eq. 5.8. It is convenient to introduce the *dual* electromagnetic field $F^*_{\mu\nu}$ defined by

$$F^{*\mu\nu} = \varepsilon^{\mu\nu\sigma\rho}F_{\sigma\rho}\,. \tag{5.16}$$

The Eq. 5.14 is then

$$F_{\mu\nu}F^{*\mu\nu} = 0. \tag{5.14$'$}$$

This quantity is one of the invariants of any electromagnetic field and Eq. 5.14 indicates that the vectors **E** and **B** are orthogonal. The second property, Eq. 5.15, shows the orthogonality of **E** and **B** to the vector $(\mathbf{x} - \mathbf{z})$. Thus, these two properties are quite general for the total field $F^{\mu\nu}$ of the particle.

The field far away from the world line, namely the radiation field which depends on the acceleration $\ddot{z}$ and varies as $1/R$ with the distance, has additional special and important properties.

The radiation field is given by[1]

$$F^{\mu\nu}(\ddot{z}) = \left(\frac{e}{4\pi}\right)\frac{1}{R^3}\left[R(x - z)^\mu\ddot{z}^\nu - (x - z)^\mu\dot{z}^\nu Q - (\mu \leftrightarrow \nu)\right]\Big|_{s=s_0} \tag{5.17}$$

where $(\mu \leftrightarrow \nu)$ means two additional terms with μ and ν interchanged and R and Q are given by Eq. 5.7.

We can now verify from Eq. 5.17 that the radiation field has the following properties in addition to Eqs. 5.14 and 5.15.

$$F_{\mu\nu}(\ddot{z})F^{\mu\nu}(\ddot{z}) = 0 \tag{5.18}$$

and

$$F_{\mu\nu}(\ddot{z})(x - z)^\nu = 0, \tag{5.19}$$

---

[1] In Section 5 we shall give a covariant definition of the radiation field and a close expression for it on the light cone. See Eqs. 5.67 and 5.71.

which follow immediately, if we note the relations

$$(x - z(s_0))^2 = 0, \qquad \dot{z}^2 = 1 \quad \text{and} \quad \dot{z}^\sigma \ddot{z}_\sigma = 0. \tag{5.20}$$

The Eq. 5.18 expresses the equality of the magnitudes of the vectors **E** and **B**.

Any electromagnetic field for which the two invariants (see Eq. 3.38) $F_{\mu\nu}F^{\mu\nu}$ and $F_{\mu\nu}F^{*\mu\nu}$ are identically zero as in the above case may be defined as a *pure radiation field*. It is also called a *null field*. Now we discuss more closely the physical significance of the radiation conditions.

Plane Wave Decomposition of the Radiation Field

Consider first a plane wave moving in a given direction, say x^1. The potentials A^μ depend on the time coordinate and one spatial coordinate, x^1, only. The subsidiary condition $A^\mu_{,\mu} = 0$ becomes

$$A^0_{,0} + A^1_{,1} = 0.$$

Let us again choose the gauge

$$A^0 = 0 \quad \text{and} \quad \nabla \cdot \mathbf{A} = 0.$$

Hence $A^1_{,1} = 0$ and the field equations reduce to the one-dimensional wave equation

$$\mathbf{A}_{,00} + \mathbf{A}_{,11} = 0.$$

For the A^1-component we obtain

$$A^1_{,00} = 0,$$

or

$$A^1_{,0} = \text{const}.$$

A constant in the last equation which is different from zero gives a constant electric field which does not represent a wave. We, therefore, choose

$$A^1 = 0.$$

For the A^2 and A^3 components of the field we have the equations

$$A^2_{,00} = A^2_{,11}$$

and

$$A^3_{,00} = A^3_{,11}.$$

Hence both A^2 and A^3 are functions of the difference $(x^0 - x^1)$, or

$$A^2_{,0} = -A^2_{,1} \quad \text{and} \quad A^3_{,0} = -A^3_{,1}.$$

From these relations we can obtain the non-vanishing field components which we show in matrix form.

$$F^{\mu\nu} = \begin{pmatrix} 0 & 0 & -E_2 & -E_3 \\ & 0 & -B_3 & B_2 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}, \quad F^{*\mu\nu} = \begin{pmatrix} 0 & 0 & B_2 & B_3 \\ & 0 & E_3 & E_2 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}$$

The field invariants are now easily calculated. We have

$$\mathbf{E} \cdot \mathbf{B} = A^2_{,0} A^3_{,1} - A^3_{,0} A^2_{,1} = 0$$

and

$$\mathbf{E}^2 - \mathbf{B}^2 = 0.$$

Thus, for a single plane wave, we have the well-known properties

$$\mathbf{E} \perp \mathbf{B} \perp \mathbf{n}$$

and

$$|\mathbf{E}| = |\mathbf{B}|$$

where **n** is the direction of propagation. However, these relations are still not sufficient to fix **E** and **B** uniquely. At a given point **x** the vectors **E**, **B** can still change as a function of time in a plane perpendicular to **n**. Thus, plane waves may still be distinguished, for a given **n**, E and B, by the direction of the vector. This directional property has observable consequences. One can therefore introduce a so-called *polarization degree of freedom* of the radiation field which then describes this additional directional property of the field. Since **E** and **B** vary in a plane, two numbers are sufficient to describe the polarization of the radiation; in other words the radiation field has *two degrees of polarization*. Still in other words we shall refer to the *transversal character* of the electromagnetic radiation field.

We describe the polarization covariantly as follows.

Consider the monochromatic plane wave

$$A^\mu(x) = C^\mu e^{-ikx} \tag{5.21}$$

with $k^2 = 0$ which is a solution of the wave equation $\Box^2 A^\mu = 0$, where C^μ is a constant vector. The subsidiary condition $A^\mu{}_{,\mu} = 0$ implies

$$k^\mu C_\mu = 0, \tag{5.22}$$

i.e., the amplitude C^μ is orthogonal to the wave vector k_μ. It is convenient to choose a k-dependent basis in the Minkowski space, i.e., four 4-vectors. (See Part I, Eq. 1.14.)

$$e^{(\lambda)}(\lambda = 0, 1, 2, 3)$$

with components $e^{(\lambda)}$ such that the components of k^μ are always

$$k^\mu e^{(\lambda)}_\mu = (k^0, k^0, 0, 0).$$

The amplitude C^μ can be expanded in the form

$$C_\mu = \sum_\lambda b^{(\lambda)} e_\mu^{(\lambda)}.$$

Now the condition, Eq. 5.22, gives

$$\sum_\lambda b^{(\lambda)} k^\mu e_\mu^{(\lambda)} = k^0(b^{(0)} - b^{(1)}) = 0$$

or

$$b^{(0)} = b^{(1)}.$$

The magnitude of $b^{(\lambda)}$ is then

$$b^2 = -(b^{(2)^2} + b^{(3)^2}).$$

The electromagnetic field corresponding to Eq. 5.21 is

$$F^{\mu\nu} = i(C^\mu k^\nu - C^\nu k^\mu)e^{-ikx}. \tag{5.23}$$

This equation taken in the basis $e_\mu^{(\lambda)}$ clearly shows the transverse character of the field: The components of **E** and **B** in the direction of **k** vanish:

$$F^{01} = 0, \qquad F^{23} = 0.$$

Equation 5.23 is a special case of a more general and useful theorem. Any null field can be written in the form

$$F^{\mu\nu} = l^\mu(x)\xi^\nu(x) - l^\nu(x)\xi^\mu(x) \tag{5.24}$$

where the vector fields $l^\mu(x)$ and $\xi^\nu(x)$ are such that

$$l^2 = 0$$

and

$$l^\mu \xi_\mu = 0; \tag{5.25}$$

hence ξ_μ is a space-like vector, $\xi^2 < 0$. With Eq. 5.24 the two-conditions

$$F^{\mu\nu}F_{\mu\nu} = 0$$

and

$$F^{\mu\nu}F^*_{\mu\nu} = 0$$

are satisfied.

For an arbitrary field satisfying the field equation $\Box^2 A^\mu = 0$ and the subsidiary condition $A^\mu_{,\mu} = 0$ we make the Fourier decomposition

$$A^\mu(x) = (2\pi)^{-4} \int a^\mu(k)\, e^{-ikx}\, dk. \tag{5.26}$$

From the reality of the field $A^\mu(x)$, the complex Fourier components $a^\mu(k)$ satisfy

$$a_\mu(-k) = a^*_\mu(k). \tag{5.27}$$

If we insert Eq. 5.26 into the field equation and the subsidiary condition, we obtain the conditions

$$k^2 = 0$$

and

$$a_\mu(k)k^\mu = 0. \tag{5.27}$$

The Fourier decomposition of the field tensor $F^{\mu\nu}$ is therefore

$$F^{\mu\nu}(x) = i\,(2\pi)^{-4} \int i(a^\mu(k)k^\nu - a^\nu(k)k^\mu)\, e^{-ikx}\, dk. \tag{5.28}$$

Energy and Momentum of the Radiation Field

The energy momentum tensor

$$T^{\mu\nu} = F^{\mu\sigma}F^\nu_\sigma - g^{\mu\nu}L$$

with

$$L = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}$$

can also be written in the symmetric form[1]

$$T^{\mu\nu} = \tfrac{1}{2}(F^{\mu\sigma}F_\sigma{}^\nu + F^{*\mu\sigma}F^{*\nu}_\sigma). \tag{5.29}$$

For a single plane wave with $F^{\mu\nu}$ and $F^{*\mu\nu}$ evaluated above we get

$$T^{\mu\nu} = \begin{pmatrix} E^2 & E^2n_1 & E^2n_2 & E^2n_3 & \\ & E^2 & 0 & 0 & \\ & & -E^2 & -E_2E_3 & -B_2B_3 \\ & & & -E^2 & \end{pmatrix}$$

where $\mathbf{n}$ (n_1, n_2, n_3) is the direction of propagation of the plane wave. We see by comparison with Eq. 3.95 that the Poynting vector is

$$S = cE^2\mathbf{n}$$

and the momentum density of field

$$G = \frac{1}{c^2}\, E^2\mathbf{n}.$$

In covariant notation the energy momentum tensor for the monochromatic wave, Eq. 5.21, is given by

$$T^{\mu\nu} = C^2k^\mu k^\nu\, e^{-2ikx} \tag{5.31}$$

---

[1] For two tensors $F^{\mu\nu}$ and $G^{\mu\nu}$ with their duals $F^*{}_{\mu\nu}$ and $G^*{}_{\mu\nu}$ we have the general formula

$$\tfrac{1}{2}g^{\mu\nu}F^{\sigma\varrho}G_{\sigma\varrho} = F^{\mu\sigma}G_\sigma{}^\nu - G^{*\mu\sigma}F_\sigma{}^{*\nu}. \tag{5.30}$$

In Eq. 5.29 we have used a special case of this formula with $G^{\mu\nu} = F^{\mu\nu}$.

which in view of the light-like momentum vector, $k^2 = 0$, satisfies the relations

$$T^{\mu\nu} k_\nu = 0,$$

$$T^\mu{}_\mu = 0 \tag{5.32}$$

and

$$T^{\mu\nu} T_{\mu\nu} = 0.$$

In general, for any *null field* given by Eq. 5.24, the energy momentum tensor can be written in the form

$$T^{\mu\nu}(x) = \xi^2 l^\mu(x) l^\nu(x) \tag{5.33}$$

with $l^2 = 0$. Clearly,

$$T^{\mu\nu} l_\nu = 0$$

and also the two last properties in Eq. 5.32 are satisfied. The matrix T has, thus, a light-like eigenvector with the eigenvalue zero.

For a general radiation field given by its Fourier decomposition, Eq. 5.26, we obtain

$$\begin{aligned} T^{\mu\nu} = -(2\pi)^{-8} \int e^{-ix(k+k')}\, dk dk' \{ & [a^\mu(k) a_\sigma(k') k^\sigma k'^\nu \\ & - a^\mu(k) a^\nu(k') k^\sigma k'_\sigma - a^\sigma(k) a_\sigma(k') k^\mu k'^\nu + a^\sigma(k) a^\nu(k') k^\mu k'_\sigma] \\ & + \tfrac{1}{2} g^{\mu\nu} [a^\sigma(k) a_\sigma(k') k^\rho k'_\rho - a^\sigma(k) a^\rho(k') k_\sigma k'_\rho] \} \end{aligned} \tag{5.34}$$

In the evaluation of the energy-momentum four vector

$$P^\nu = \int d\sigma_\mu T^{\mu\nu} = \int d^3x T^{0\nu}$$

The x-integration can be carried out giving $\delta(\mathbf{k} + \mathbf{k}')$. Then $\mathbf{k} + \mathbf{k}' = 0$ implies $k'_0 = \pm k_0$, or

$$k = -k'.$$

Thus, we find

$$P^0 = (2\pi)^{-3} \sum_{\sigma=2,3} \int d\mathbf{k}\, a^\sigma(k) a^*_\sigma(k) \mathbf{k}^2, \tag{5.35}$$

where in the basis $e^{(\lambda)}_\mu$ introduced above we have to sum only over the transverse indices of $a^\mu(k)$.

Equation 5.35 expresses the total field energy as a sum over plane waves each with two transverse degrees of polarization.

We remark that the Fourier decomposition, Eq. 5.26, can also be written as

$$A^\mu(k) = (2\pi)^{-3} \int dk e^{-ikx}\, \delta(k^2) A^\mu(k)$$

where the coefficients $A^\mu(k)$ are different than the $a^\mu(k)$ in Eq. 5.26. The same results can be obtained by using this expansion.

3. *Canonical Formalism for the Radiation Field in Terms of the Transverse Vector Potential*

Finally we make use of the transverse character of the electromagnetic field to give a canonical form of the field equations which eliminates the redundant components from the beginning.

Any vector **A** can be split in the form

$$\mathbf{A} = \mathbf{A}^T + \mathbf{A}^L \tag{5.36}$$

where, by definition, the transverse component satisfies

$$\nabla \cdot \mathbf{A}^T = 0 \tag{5.37}$$

and the longitudinal part is soleinoidal

$$\nabla \times \mathbf{A}^L = 0, \tag{5.38}$$

and, therefore, must be the gradient of some potential

$$\mathbf{A}^L = \nabla\psi \tag{5.39}$$

The components $\mathbf{A}^T$ and $\mathbf{A}^L$ are "orthogonal" in the following sense

$$\int d^3x \mathbf{A}^T \cdot \mathbf{A}^L = -\int (\nabla \cdot \mathbf{A}^T)\psi \, d^3x = 0.$$

The subsidiary condition $A^\mu_{,\mu} = 0$ in the gauge $A^\mu(0, A)$ becomes the *radiation gauge*

$$\nabla \cdot \mathbf{A} = 0,$$

which implies that

$$\nabla \cdot \mathbf{A}^L = 0 \quad \text{or} \quad \nabla^2\psi = 0. \tag{5.40}$$

The only solution of the Laplace equation which is regular everywhere (in the absence of sources) is a constant. Hence we can put

$$\mathbf{A}^L = 0 \tag{5.41}$$

Because

$$\mathbf{E} = -\frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}$$

we can make a similar splitting for **E** into its longitudinal and transverse components and obtain

$$\mathbf{E}^L = 0$$

and

$$\nabla \cdot \mathbf{E}^T = 0. \tag{5.42}$$

The field equations are now

$$\Box^2 \mathbf{A}^T = \partial^0 \partial_0 \mathbf{A}^T + \partial^k \partial_k \mathbf{A}^T = 0; \tag{5.43}$$

no subsidiary condition is needed, since $\nabla \cdot \mathbf{A}^T = 0$ by definition.

From the second order Lagrange equation, Eq. 5.43, we can pass to first order canonical equations

$$\begin{aligned} \mathbf{E} &= -\partial_0 \mathbf{A}^T \\ \text{and} \qquad & \\ \partial_0 \mathbf{E} &= -(\nabla \cdot \nabla) \mathbf{A}^T. \end{aligned} \tag{5.44}$$

These Hamilton's equations are gauge invariant, because $\mathbf{A}^T$ is. Under a gauge transformation $A^\mu \to A^\mu + \Lambda'^\mu$ the new potential

$$\mathbf{A}' = \mathbf{A}^{T'} + A^{L'}$$

is related to the original one by

$$\mathbf{A}' = A^T + \nabla \Lambda.$$

However, since $\nabla \times \nabla \Lambda = 0$, we have

$$\mathbf{A}^{T'} = \mathbf{A}^T \quad \text{and} \quad A_0' = \partial_0 \Lambda.$$

In Eqs. 5.44 $\mathbf{E}$ and $\mathbf{A}^T$ may be considered as the conjugate set of variables. This is an important point for quantization.

The canonical equations, Eq. 5.44, do not look manifestly covariant; $\mathbf{E}$ and in particular $\mathbf{A}^T$ do not have simple transformation properties. The equations are, however, relativistically invariant.

The essential point in the derivation of the canonical Eqs. 5.44 was the use of the radiation gauge $\nabla \cdot \mathbf{A} = 0$. In order to have the same canonical equations and the same physical interpretation, we must impose $\nabla \cdot \mathrm{A} = 0$ in every frame. This is possible only if we make, together with every Lorentz transformation, also a suitable gauge transformation to ensure the validity of the radiation gauge. Let

$$A_\mu'(x') = L_\mu^{\ \nu} A_\nu(x)$$

be the transformed field; the field in the new frame, however, will be taken to be

$$\tilde{A}_\mu(x') = L_\mu^{\ \nu} A_\nu(x) + \Lambda_{,\mu}(x),$$

where we choose the gauge Λ in such a way that

$$\nabla \cdot \tilde{\mathbf{A}} = \partial \tilde{A}_k / \partial x_k' = \frac{\partial}{\partial x_j} (L_k^{\ \nu} A_\nu(x)) \frac{\partial x_j}{\partial x_k'} + \Lambda_{,k}{}^j \frac{\partial x_j}{\partial x_k'} = 0.$$

Hence we find

$$\Lambda_{k,}{}^j = -L_k^{\ \nu} A_\nu{}^j.$$

In mathematical language the invariance of the canonical Eqs. 5.44 involve more general representations of the Lorentz group; $\mathbf{E}$ and $\mathbf{A}^T$ are geometrical objects transforming according to this, what we may call, the *gauge dependent representations* of the Lorentz group. If we denote these representations by $D(L, \Lambda)$ the group property reads now

$$D(L_2\Lambda_2)D(L_1\Lambda_1) = D(L, \Lambda),$$

the identity element being $D(I, 0)$.

Quite generally, if a theory admits more than one invariance principle, or group of transformations T_i, we can use *joint representations* $D(T_i)$. For example, every time we make a Lorentz transformation we can also simultaneously transform our physical quantities under T_i. This may lead to simplified theories as was the case for the canonical treatment of the electromagnetic field.

4. *Energy and Momentum Radiated*

In this section we discuss the energy balance between the particle and its field and energy and momentum lost by radiation to infinity.

In the wave zone the invariant $F^{\mu\nu}F_{\mu\nu}$ vanishes and the energy-momentum tensor is given by the simple expression

$$T^{\mu\nu}(\ddot{z}) = F^{\mu\sigma}(\ddot{z})F_{\sigma}{}^{\nu}(\ddot{z}),$$

where $F^{\mu\nu}(\ddot{z})$ is the far field depending on the acceleration and given by Eq. 5.17. A straightforward calculation gives

$$T^{\mu\nu}(\ddot{z}) = (e/4\pi)^2(1/R^6)\{[R(x-z)^{\mu}\ddot{z}^{\sigma} - Q(x-z)^{\mu}\dot{z}^{\sigma} - (\mu \leftrightarrow \sigma)]$$
$$\times [R(x-z)_{\sigma}\ddot{z}^{\nu} - Q(x-z)_{\sigma}\dot{z}^{\nu} - (\sigma \leftrightarrow \mu)]\}\big|_{s=s_0}$$

where again $(\mu \leftrightarrow \nu)$ means terms with μ and ν interchanged; R and Q have been defined by Eq. 5.7.

Noting that

$$(x - z(s_0))^2 = 0$$

and

$$\dot{z}(s_0^{\sigma})\ddot{z}_{\sigma}(s_0) = 0$$

we find

$$T^{\mu\nu}(\ddot{z}) = \left(\frac{e}{4\pi}\right)^2 \frac{(x-z)^{\mu}(x-z)^{\nu}}{R^4}\left(\ddot{z}^2 + \frac{Q^2}{R^2}\right)\bigg|_{s=s_0}. \tag{5.46}$$

We show first that $T^{\mu\nu}(\ddot{z})$ is positive definite as it should be. The acceleration $\ddot{z}$ is a space-like vector (in the rest frame of the particle

$\ddot{z}$: $(0, \ddot{\mathbf{z}})$), hence $\ddot{z}^2 < 0$. Furthermore, we decompose $(x - z)$ into a time-like and an orthogonal space-like vector (see Eq. (1.9)). Let

$$(x - z)_{s_0} = R\dot{z} + y \tag{5.47}$$

where

$$y^2 < 0 \quad \text{and} \quad \dot{z}y = 0.$$

Indeed, we obtain

$$(x - z)^\sigma \dot{z}_\sigma = R > 0$$

and

$$Q = (x - z)^\sigma \ddot{z}_\sigma = y^\sigma \ddot{z}_\sigma .$$

If we form

$$0 = (x - z(s_0))^2 = R\dot{z}^2 + y^2 = R^2 + y^2$$

we find

$$y^2 = -R^2.$$

Now from Schwarz inequality (Chapter I),

$$Q^2 = (y^\sigma \ddot{z}_\sigma)^2 \leqq \ddot{z}^2 y^2 = R^2 |\ddot{z}|^2$$

or

$$Q^2/R^2 \leqq |\ddot{z}|^2.$$

This equation shows that $T^{\mu\nu}(\ddot{z})$ is indeed positive definite. Moreover, $T^{\mu\nu}$ is the greater as Q is smaller. We obtain the maximum of $T^{\mu\nu}$ where Q vanishes, i.e., when

$$(x - z)^\sigma \ddot{z}_\sigma = 0;$$

in the rest frame, $(\mathbf{x} - \mathbf{z}) \cdot \ddot{\mathbf{z}} = 0$; i.e., in a direction $(\mathbf{x} - \mathbf{z})$ perpendicular to the acceleration the energy momentum tensor has its maximum.

From the conservation law

$$T^{\mu\nu}{}_{,\mu} = -\tau^{\mu\nu}{}_{,\mu},$$

where $\tau^{\mu\nu}$ is the matter tensor of the particle (see Eq. 4.10), we now evaluate the energy and momentum given from the particle to the field in some space-time region v *including the particle world line*:

$$\Delta P^\mu = -\int_v \tau^{\mu\nu}{}_{,\nu}\, dx = +\int_v dx\, T^{\mu\nu}{}_{,\nu} . \tag{5.48}$$

The field produced by the event $z(s_0)$ propagates along the forward light-cone (outgoing spherical waves with velocity c). The field produced by the event $z(s_0 + ds)$ propagates in a similar light cone, so that the field associated with the arc AB of the world line of the particle is confined to the region between the two future light cones. (Fig. 10). To evaluate the energy and momentum radiated, we take a large three-dimensional surface S and calculate the energy and momentum flux across it. We take, thus, as the volume v in Eq. 5.48 the region surrounded by the two light cones and the surface S.

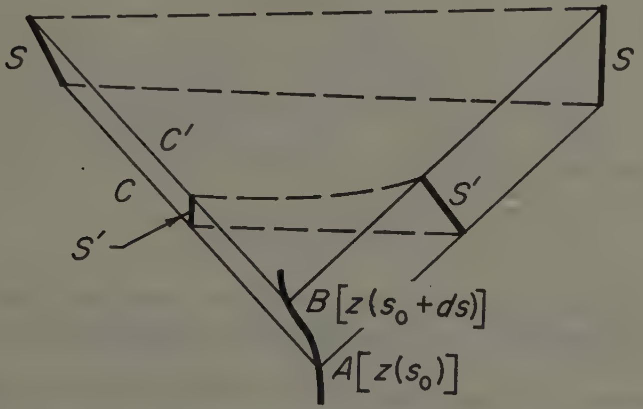


Fig. 10. The radiation produced by the particle along its trajectory from *A* to *B* is confined to the region between the two forward light cones *C* and *C′*.

We transform the volume integral, Eq. 5.48, into a surface integral

$$\Delta P^{\mu} = \int_{\text{Boundary}} d\sigma_{\nu} T^{\mu\nu} \tag{5.49}$$

Part of the boundary, namely the surfaces of the cones, C and C', does not contribute to the integral, Eq. 5.49. For $T^{\mu\nu}$ is proportional to $(x - z)^{\nu}$ and along C and C', $d\sigma_{\nu}$ is orthogonal to $(x - z)^{\nu}$. Physically, this means that the energy, of course, does not cross the two cones. Therefore, the outward radiation of energy is given by

$$\Delta P^{\mu} = \int_{S} d\sigma_{\nu} T^{\mu\nu}, \tag{5.50}$$

as we have anticipated. We have written ΔP^{μ} to show that this radiation is associated with a small segment of the motion of the particle.

Since the choice of the surface S is arbitrary and, in particular, different observers may choose different surfaces, we must first show that ΔP^{μ} defined by Eq. 5.50 is independent of the choice of the surface S. This will establish the four-vector character of ΔP^{μ} which is essential.

For this purpose consider another surface S' and the region v' of the space-time surrounded by S, S' and the two-light cones, C and C'. Since this region of space-time does not contain the world line of the particle, we have the conservation law

$$\int_v T^{\mu\nu}{}_{,\nu}\, dx = 0$$

or

$$\int_S d\sigma_\nu T^{\mu\nu} - \int_{S'} d\sigma_\nu T^{\mu\nu} + \int_C d\sigma_\nu T^{\mu\nu} \int_{C'} d\sigma_\nu T^{\mu\nu} = 0.$$

Again the last two integrals are zero[1] and we have

$$\int_S d\sigma_\nu T^{\mu\nu} = \int_{S'} d\sigma_\nu T^{\mu\nu}, \tag{5.51}$$

i.e., the energy momentum crossing the surface S' will also cross S. ΔP^μ being independent of the choice of the three-dimensional surface S, is a four-vector.

The surface S is space-like; its normal is in the direction of the space-component of $(x - z)$ according to the decomposition

$$(x - z) = R\dot{z} + y.$$

Thus, we write the surface element on S as

$$d\sigma_\mu = y_\mu\, dsRd\Omega;$$

In the rest frame S can be taken to be a sphere of radius R; $d\Omega$ is the solid angle. We have then

$$\Delta P^\mu = -\int dsRd\Omega y_\mu T^{\mu\nu},$$

or with Eq. 5.46,

$$\Delta P^\mu = -(e/4\pi)^2\, ds \int d\Omega \frac{1}{R^3}\, y_\nu (R\dot{z}^\mu + y^\mu)(R\dot{z}^\nu + y^\nu)\left[\ddot{z}^2 + \frac{Q^2}{R^2}\right]\Bigg|_{so}.$$

This equation simplifies by noting that

$$y_\nu \dot{z}^\nu = 0 \quad \text{and} \quad y_\nu y^\nu = -R^2.$$

$$\Delta P^\mu = -(e/4\pi)^2\, ds \int \frac{d\Omega}{R} (R\dot{z}^\mu + y^\mu)\left(\ddot{z}^2 + \frac{Q^2}{R^2}\right)$$

Now the integral $\int d\Omega y^\mu \cdots$ is zero, since y is orthogonal to the surface. We further put

$$Q^2 = (y^\sigma \ddot{z}_\sigma)^2 = y^2 \ddot{z}^2 \cos^2\theta = -R^2 \ddot{z}^2 \cos^2\theta,$$

---

[1] These integrals are not zero if we take the total field rather than the far field. But the contributions, of course, vanish at large distances.

where θ is just the azimuth of $d\Omega$. Then

$$\Delta P^\mu = -(e/4\pi)^2\, ds \int d\Omega \dot{z}^\mu \ddot{z}^2 (1 - \cos^2\theta)$$

or

$$\frac{\Delta P^\mu}{ds} = -\frac{e^2}{4\pi}\frac{2}{3}\dot{z}^\mu \ddot{z}^2 \tag{5.52}$$

This is the general covariant formula for the energy and momentum four vector radiated per unit proper time.

Writing

$$\dot{z} = \frac{dt}{ds}(c,\mathbf{v})$$

and

$$\ddot{z} = \frac{dt}{ds}\left[\frac{dt}{ds}(0,\dot{\mathbf{v}}) + (c,\mathbf{v})\frac{d}{dt}\left(\frac{dt}{ds}\right)\right]; \qquad \dot{\mathbf{v}} = \frac{d\mathbf{v}}{dt}.$$

With $\dfrac{dt}{ds} = \gamma/c$, $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$ we get

$$\ddot{z}^2 = -\frac{\gamma^4}{c^4}\Bigg(\dot{\mathbf{v}}^2 + 2\left(\frac{\gamma}{c}\mathbf{v}\cdot\dot{\mathbf{v}}\right)^2\Bigg) + \frac{\gamma^4}{c^4}(\mathbf{v}\cdot\dot{\mathbf{v}})^2 v^2\Bigg) + c^2\left(\frac{\gamma}{c}\right)^8(\mathbf{v}\cdot\dot{\mathbf{v}})^2,$$

hence

$$\frac{\Delta P^\mu}{dt} \cong (e^2/4\pi)\tfrac{2}{3}(\dot{z}^\mu)(\gamma/c)^4\left[\dot{\mathbf{v}}^2 + \left(\frac{\gamma}{c}\mathbf{v}\cdot\dot{\mathbf{v}}\right)^2\right] \tag{5.53}$$

which, in the nonrelativistic limit, becomes

$$\frac{\Delta P^0}{dt} = (e^2/4\pi)\frac{2}{3}\frac{\dot{\mathbf{v}}^2}{c^3} \tag{5.54}$$

$$\frac{\Delta \mathbf{P}}{dt} = (e^2/4\pi)\frac{2}{3}\frac{\dot{\mathbf{v}}^2}{c^4}\mathbf{v} \tag{5.55}$$

For further use we also evaluate $\dddot{z}$:

$$\dddot{z} = (0,\dot{\mathbf{v}})\frac{\gamma}{c}\frac{d}{dt}\left(\frac{\gamma^2}{c^2}\right) + (0,\ddot{\mathbf{v}})\frac{\gamma^2}{c^2} + (0,\dot{\mathbf{v}})\frac{\gamma^2}{c^2}\frac{d}{dt}\left(\frac{\gamma}{c}\right) + (c,\mathbf{v})\left[\left(\frac{d}{dt}\left(\frac{\gamma}{c}\right)\right)^2 + \left(\frac{\gamma}{c}\right)^2\frac{d^2}{dt^2}\left(\frac{\gamma}{c}\right)\right],$$

with

$$\frac{d}{dt}\left(\frac{\gamma}{c}\right) = (\gamma/c)^3\,\mathbf{v}\cdot\dot{\mathbf{v}}$$

$$\frac{d}{dt}(\gamma^2/c^2) = 2(\gamma/c)^4\,\mathbf{v}\cdot\dot{\mathbf{v}}$$

$$\frac{d^2}{dt^2}(\gamma/c) = \left(\frac{\gamma}{c}\right)^3\left[\dot{v}^2 + \mathbf{v}\cdot\ddot{\mathbf{v}} + 3\left(\frac{\gamma}{c}\right)^2(\mathbf{v}\cdot\dot{\mathbf{v}})^2\right]$$

For a finite interval of the world line AB the total radiation is given by

$$P^\mu = -\int_A^B (e^2/4\pi)\tfrac{2}{3}\dot{z}^\mu\ddot{z}^2 \tag{5.56}$$

This formula assumes that the trajectory $z^\mu(s)$ is given, then P^μ may be evaluated. However, strictly speaking, $z^\mu(s)$ cannot be given a priori, because due to radiation lost the particle must change its acceleration.

5. *Radiation Reaction*

(A) Energy Balance

The radiation field of the particle changes the total field which accelerates the particle. Therefore, the motion of the particle changes. This is a type of "feedback" process. Another way of expressing this process is that with the act of radiation (or energy loss) the particle has to de-accelerate simply by energy balance.

In the light of this phenomenon the equations of the motion of a particle discussed in Part I, Chapter II, are only approximately correct, for we have treated the external fields and therefore the forces on the particle as given. One way to include the radiation reaction into the equation of motion is to consider the energy loss caused by the radiation as the result of an additional (fictitious) force on the particle, the so-called *reactive force* or *self force*, K_{react}. The Minkowski Eq. 2.10 will now be generalized into

$$\dot{p} = (m_0 c\dot{u}) = K_{\text{ext}} + K_{\text{react}}\,. \tag{5.57}$$

It is instructive to start the discussion with a simplified and heuristic derivation of K_{react} in electrodynamics based on the idea of energy balance.[1]

We consider a one-dimensional radiating motion. The energy lost by the particle in a time interval T must be equal to the work done by the self-force. In the non-relativistic limit we have then

$$\frac{1}{T}\int_t^{t+T} \frac{2}{3}\frac{e^2}{4\pi c^3}\,\ddot{x}^2\,dt = -\frac{1}{T}\int_t^{t+T} K_{\text{react}}\,\dot{x}\,dt. \tag{5.58}$$

---

[1] This derivation goes back to Planck, Abraham and Lorentz. For a historical review and a long list of references for the subject matter of this section see T. Erber, *Fortschritte der Physik*, 9, (1961), 343.

If we integrate by parts the left-hand side and choose a T such that $\ddot{x}$ vanishes at both ends of the integral, we get

$$\int_t^{t+T} \left(\frac{2}{3} \frac{e^2}{4\pi c^3} \dddot{x} - K_{\text{react}} \right) \dot{x} \, dt = 0.$$

One possible, but not necessarily the only, solution of this equation is obtained by putting the integrand equal to zero. Hence we find

$$K_{\text{react}} = \frac{2}{3} \frac{e^2}{4\pi c^3} \dddot{x}. \tag{5.59}$$

The assumption made above that $\ddot{x}$ vanishes at t and $t + T$ is certainly true for oscillatory motion, but we expect that a result like Eq. 5.59 should be quite generally true.

The form of the self force is rather unusual; it is proportional to the third derivative of the position $x(t)$. Indeed, the radiation reaction is an unusual phenomenon. Accordingly, the equation of motion is a third order differential equation

$$m\ddot{x} = F_{\text{ext}}(\dot{x}) + \tfrac{2}{3}(e^2/4\pi c^3)\dddot{x}, \tag{5.60}$$

and we need *three* initial conditions to specify a unique solution. We shall discuss the solutions of Eq. 5.60 in the next section.

If we use the relativistic formula, Eq. 5.53, for the radiation energy, the form of the self-force is

$$K_{\text{react}} = \tfrac{2}{3}(e^2/4\pi c^3)\gamma^2 \left[\ddot{\mathbf{v}} + \left(\frac{\gamma}{c}\right)^2 (\mathbf{v}\cdot\ddot{\mathbf{v}})\mathbf{v} + 3\left(\frac{\gamma}{c}\right)^2 (\mathbf{v}\cdot\dot{\mathbf{v}})\dot{\mathbf{v}} + 3\left(\frac{\gamma}{c}\right)^4 (\mathbf{v}\cdot\dot{\mathbf{v}})^2\mathbf{v} \right] \tag{5.61}$$

or, in covariant form

$$K_{\text{react}} = -\frac{2}{3} \frac{e^2}{4\pi c^3} [\dddot{z}^{\mu} + \dot{z}^{\mu}\ddot{z}^2] \tag{5.61$'$}$$

where the dot is now differentiation with respect to the proper time s. Equation 5.61 is, indeed, the unique covariant form of Eq. 5.59.[1]

The above treatment is, as we had said, heuristic and phenomenological. Even in the derivation of the radiation in Section 3 we have assumed the motion or the world line of the particle as given. The correct treatment of radiation and radiation reaction must consider both the particle and the field simultaneously. This will give us also an insight into the mechanism of the radiation reaction.

[1] M. von Laue, *Ann. Phys.*, *28*, (1909), 436.

(B) Interacting Fields and Particles

Let $F^{\mu\nu}_{\text{in}}(x)$ be the given free external field which starts the motion and $F^{\mu\nu}$ the self field of the particle. The total Lagrangian of a single charged particle interacting with these fields is

$$L = m_0 cv^2 + \tfrac{1}{4}(F^{\mu\nu} + F^{\mu\nu}_{\text{in}})(F_{\mu\nu} + F^{\text{in}}_{\mu\nu}) + \frac{e}{c}\int ds\dot{z}_\mu \rho(x - z)(A^\mu + A^\mu_{\text{in}}) \tag{5.62}$$

where in the last interaction term A^μ and A^μ_{in} are the potentials corresponding to $F^{\mu\nu}$ and $F^{\mu\nu}_{\text{in}}$, and in the expression for current density, Eq. 4.14, we have replaced $\delta(x - z)$ by an *invariant density* of the charge $\rho(x - z)$ for reasons which will be explained below. At any rate, in all the following expressions $\rho(x - z)$ will be replaced by $\delta(x - z)$ to get the results for a point particle.

If we now vary in Eq. 5.62 A_μ and A^μ_{in} independently, we obtain (a) the equation of the motion for the particle

$$m_0 c\ddot{z}_\mu = \frac{e}{c}\int ds\dot{z}^\nu \rho(x - z)[F^{\text{ret}}_{\mu\nu} + F^{\text{in}}_{\mu\nu}] \tag{5.63}$$

(b) the equation of the motion for the self-field

$$F^{\mu\nu}_{\text{ret},\nu} = -\frac{e}{c}\int ds\rho(x - z)\dot{z}^\mu \tag{5.64}$$

and (c) the equation of the motion for the external field

$$F^{\mu\nu}_{\text{in},\nu} = -\frac{e}{c}\int ds\rho(x - z)\dot{z}^\mu \approx 0. \tag{5.65}$$

The fields $F^{\mu\nu}$ and $F^{\mu\nu}_{\text{in}}$ satisfy different boundary conditions, in particular the external field usually is so strong that the inhomogeneous term in Eq. 5.65 can be neglected, so that we need only to consider the coupled Eqs. 5.63 and 5.64.

In Eq. 5.63 the second term on the right-hand side is the usual external force. The first term represents the *self-force*. For point particles we see that to evaluate the self-force we need the self-field on the world line itself. It is easy to see that the self-field evaluated on the world-line itself is infinite. For example, we may consider the retarded field of the particle

$$A^{\text{ret}}_\mu = \frac{1}{c}\int dx' D^{\text{ret}}(x - x')j_\mu(x')\, dx'$$

or the corresponding $F^{\mu\nu}_{\text{ret}}$. Since the function $D^{\text{ret}}(x - z)$ has a δ-function form, the field $F^{\mu\nu}_{\text{ret}}$ evaluated at $x = z$ is infinite; similarly $F^{\mu\nu}_{\text{ret}}$.

This may seem to contradict the earlier phenomenological result where we have found a finite and reasonable expression for the self force. When we look closer, however, we see that the two procedures are different. Whereas in the argument based on the energy balance we have only used the radiation on the far field of the particle, the self-force in Eq. 5.63 contains the *total* field produced by the point particle. It is thus clear that even for a point particle the self-force has a finite, physically meaningful part associated with the radiation field, and an infinite part associated with the near field and the mass point character of the particle.

(C) Finite Part of Self-Force

It is indeed possible to extract, in a covariant manner, the finite part of the self-force associated with the radiation field and see the nature of the infinite part. For this purpose we write[1]

$$F_{\mu\nu}^{\text{ret}} = \tfrac{1}{2}F_{\mu\nu}^{\text{rad}} + \tfrac{1}{2}(F_{\mu\nu}^{\text{ret}} + F_{\mu\nu}^{\text{adv}}) \tag{5.66}$$

where we have defined the radiation field by

$$F_{\mu\nu}^{\text{rad}} = F_{\mu\nu}^{\text{ret}} - F_{\mu\nu}^{\text{adv}} \tag{5.67}$$

or, equivalently,

$$F_{\mu\nu}^{\text{rad}} = F_{\mu\nu}^{\text{out}} - F_{\mu\nu}^{\text{in}}. \tag{5.68}$$

It will be calculated from the potential

$$A_{\mu}^{\text{rad}} = A_{\mu}^{\text{out}} - A_{\mu}^{\text{in}}$$

or

$$A_{\mu}^{\text{rad}} = \frac{1}{c}\int (D^{\text{ret}} - D^{\text{adv}})\, j_\mu(x')\, dx' = \frac{1}{c}\int D(x - x')\, j_\mu(x')\, dx' \tag{5.68$'$}$$

and represents the total change in the external field between $t = -\infty (F_{\text{in}})$ and $t = +\infty (F_{\text{out}})$.

It will now be shown that the self-force corresponding to the term $\tfrac{1}{2}F_{\mu\nu}^{\text{rad}}$ in Eq. 5.66 is finite and agrees with the result, Eq. 5.61.[2] The infinity is therefore due to the term $\tfrac{1}{2}(F_{\mu\nu}^{\text{ret}} + F_{\mu\nu}^{\text{adv}})$ in Eq. 5.66. We first calculate $F_{\mu\nu}^{\text{rad}}$ by a procedure analogous to the calculation of the total field in Sect. 1 of the present Chapter.

---

[1] The splitting in Eq. 5.66 is not unique; more generally one can put

$$F_{\mu\nu}{}^{\text{ret}} = k(F^{\text{ret}} - F^{\text{adv}}) + (1 - k)F^{\text{ret}} + kF^{\text{adv}}.$$

This, however, does not change the essential result to be derived.

[2] P. A. M. Dirac, *Proc. Roy. Soc., A 167,* (1938), 148; P. A. M. Dirac, *Ann. Institut Poincaré 9,* (1938), 13.

From Eq. 5.68 we get

$$F^{\text{rad}}_{\mu\nu}(x) = e \int ds D(x - z(s)) \frac{d}{ds} \left[\frac{\dot{z}_\mu (x - z)^\nu}{(x - z)^\sigma \dot{z}_\sigma} - (\mu \leftrightarrow \nu) \right] \tag{5.69}$$

where, this is the difference from the previous calculation, the homogeneous Green's function is given by

$$D(x) = \frac{1}{2\pi} \delta(x^2) \varepsilon(x^0).$$

Actually, we must evaluate, for a point particle, $F^{\text{rad}}_{\mu\nu}$ at $x = z(s)$. Let $z(s')$ be a fixed point on the world line and set

$$s = s' + u.$$

Then

$$z(s) - z(s') = u\dot{z} + \frac{u^2}{2} \ddot{z} + \frac{u^3}{6} \dddot{z} + \cdots$$

and

$$\dot{z}(s) = \dot{z} + u\ddot{z} + \frac{u^2}{2} \dddot{z} + \cdots.$$

Here all derivatives $\dot{z}, \ddot{z}, \dddot{z}, \cdots$ refer to the fixed point s'. Hence

$$(z(s) - z(s'))^2 = u^2 + O(u^4)$$

and

$$(z(s') - z(s))^\sigma \dot{z}_\sigma = -u + O(u^3)$$

since $\dot{z}^2 = 1$. Furthermore, $z^0(s) - z^0(s')$ has the same sign as u ($\dot{z}$ is time-like) which implies that

$$\varepsilon(z(s') - z(s)) = -\varepsilon(u).$$

If we insert these expansions into Eq. 5.69 at $x = z(s')$ we get

$$F^{\text{rad}}_{\mu\nu}(z(s')) = \frac{e}{2\pi} \int_{-\infty}^{+\infty} du \varepsilon(u) \delta(u^2) \frac{d}{du}$$

$$\left[u\ddot{z}_\mu \dot{z}_\nu + \frac{u}{2} \dot{z}_\mu \ddot{z}_\nu + \frac{u^2}{2} \dddot{z}_\mu \dot{z}_\nu + \frac{u^2}{6} \dot{z}_\mu \dddot{z}_\nu - (\mu \leftrightarrow \nu) \right]. \tag{5.70}$$

If we use the identity[1]

$$\varepsilon(u)\delta(u^2) = -\delta'(u).$$

---

[1] To prove this identity we write

$$\delta(u^2) = \lim_{a \to 0} \delta(u^2 - a^2) = \lim_{a \to 0} \left[\frac{\delta(u - a)}{2a} + \frac{\delta(u + a)}{2a} \right], a > 0$$

then

$$\varepsilon(u)\delta(u) = \lim_{a \to 0} \left[\frac{\delta(u - a)}{2a} - \frac{\delta(u + a)}{2a} \right] = -\delta'(u).$$

We can integrate Eq. 5.70 by parts and obtain

$$F_{\mu\nu}^{\text{rad}}(z(s')) = \frac{e}{2\pi} \int du\delta(u)\, \frac{d_2}{du^2}$$

$$\left[u\ddot{z}_\mu \dot{z}_\nu + \frac{u}{2}\, \dot{z}_\mu \ddot{z}_\nu + \frac{u^2}{2}\, \dddot{z}_\mu \dot{z}_\nu + \frac{u^2}{6}\, \dot{z}_{\mu\nu}\dddot{z} - (\mu \leftrightarrow \nu)\right].$$

After differentiating with respect to u twice and doing the δ-integration, we see that only the third and the fourth terms in the bracket contributes. Also higher order terms in u give vanishing contribution and we find a covariant and simple expression for the radiation field.

$$F_{\mu\nu}^{\text{rad}}(z(s)) = \frac{2}{3}\frac{e}{2\pi}\, (\dot{z}_\mu \dddot{z}_\nu - \dot{z}_\nu \dddot{z}_\mu). \tag{5.71}$$

We can now use this expression to evaluate the corresponding self force. From Eqs. 5.66 and 5.63 we have for point particles

$$\begin{aligned} K^{\text{rad}}_{\text{self}\,u} &= \frac{e}{c} \int ds\dot{z}^\nu \delta(x - z(s))\tfrac{1}{2}F_{\mu\nu}^{\text{rad}} \qquad (5.72)\\ &= \frac{e^2}{4\pi}\, \tfrac{2}{3}(\dot{z}^\nu \dot{z}_\mu \dddot{z}_\nu - \dddot{z}_\mu). \end{aligned}$$

If we differentiate $\dot{z}^\mu \dot{z}_\mu = 1$ we get $\dot{z}^\mu \ddot{z}_\mu = 0$ and if we differentiate once more we have

$$\dot{z}_\nu \dddot{z}^\nu = -\ddot{z}^2.$$

We have finally

$$K^{\text{rad}}_{\mu\,\text{self}} = -\tfrac{2}{3}(e^2/4\pi)(\dddot{z}_\mu + \dot{z}_\mu \ddot{z}^2). \tag{5.73}$$

The notation on the left-hand side indicates the self-force on the particle due to the radiation field.

The final result, Eq. 5.73, which agrees with Eq. 5.61′, is true only for a point particle; for an extended particle one has to make a new calculation from Eq. 5.72 with $\delta(x - z)$ replaced by $\rho(x - z)$ which has to be a given function.

Equation 5.73 has another interesting aspect. If we equate K_μ^{rad} to the energy momentum per unit time given to the field

$$K^{\text{rad}}_{\mu\,\text{self}} = \frac{\Delta P_\mu}{ds},$$

then

$$\frac{\Delta P^\mu}{ds} = -\tfrac{2}{3}(e^2/4\pi)\dot{z}^\mu \ddot{z}^2 - \tfrac{2}{3}(e^2/4\pi)\dddot{z}^\mu.$$

The first term is exactly the radiation energy momentum, Eq. 5.52, which is positive definite and represents energy and momentum which is given to the field irreversibly. The second term, on the other hand, is not positive definite and must be associated with energy and momentum which is emitted and absorbed by the particle. To agree with the energy balance relation Eq. 5.58, we must now show that the integral

$$\int_t^{t+T} \dddot{z}^\mu \dot{z}_\mu \, ds = -\int_t^{t+T} \ddot{z}^2 \, ds$$

vanishes if T is such that $\ddot{z}$ vanishes at both ends of the integral, which is obviously the case. The vanishing of $\ddot{z}$ at both ends was the same assumption which led us from Eq. 5.58 to Eq. 5.59.

(D) Mass Renormalization

Now we turn to the infinite part of the self-force

$$K^\infty_{\mu\,\text{self}} = \frac{1}{2}\frac{e}{c}\int ds \dot{z}^\nu \delta(x-z)(F^{\text{ret}}_{\mu\nu} + F^{\text{adv}}_{\mu\nu}) \tag{5.74}$$

where

$$F^{\text{ret}}_{\mu\nu} + F^{\text{adv}}_{\mu\nu} = e\int ds(D^{\text{ret}} + D^{\text{adv}})\frac{d}{ds}\left[\frac{\dot{z}_\mu (x-z)\nu}{(x-z)^\sigma \dot{z}_\sigma} - (\mu \leftrightarrow \nu)\right]$$

with

$$D^{\text{ret}} = \frac{1}{2\pi}\delta((x-z)^2)\theta(x^0 - z^0) \tag{5.75}$$

and

$$D^{\text{adv}} = \frac{1}{2\pi}\delta((x-z)^2)\theta(z^0 - x^0).$$

The field, Eq. 5.74, is infinite at the point $x = z(s)$. Physically, we know that it represents the effect of the near field of the particle. In quantum mechanical language the particle carries with itself a cloud of photons which are continuously emitted and absorbed. It is, therefore, to be expected that this force contributes to the inertia or mass of the particle. The near field is always present even if there is no radiation, for example, for a uniformly moving particle. Thus, the experimentally measured mass already includes this effect. In other words, if we can show that the force in Eq. 5.74 has the form

$$K^\infty_{\mu\,\text{self}} = -\delta m c \ddot{z}_\mu \tag{5.76}$$

then the equation of motion, Eq. 5.63, can be written with Eq. 5.66 or Eq. 5.73 as

$$(m + \delta m)c\ddot{z}_\mu = \frac{e}{c}F^{\text{in}}_{\mu\nu}(z)\dot{z}^\nu + \tfrac{2}{3}(e^2/4\pi)(\dddot{z}_\mu + \dot{z}_\mu \ddot{z}^2). \tag{5.77}$$

We can then put

$$m + \delta m = m_{\text{exp}} \tag{5.78}$$

and would have a meaningful, finite and complete equation of motion for the particle. This is the form of the so-called "mass-renormalization" in classical electrodynamics. Mass and charge renormalizations play an important role in quantum field theory.

The significance of the procedure of mass-renormalization is that the original or bare mass m does not really have a physical significance; one always measures $m + \delta m = m_{\text{exp}}$. Even if we neglect, as is done most of the time, the radiation effects, we cannot neglect the effect of the near field which results in δm-term.

The procedure has also allowed us to incorporate a formally infinite expression into a finite one, m_{exp}. One would be inclined to say that the original unobservable bare mass m is also somehow infinite due to the point nature of the particle so that the sum $m + \delta m$ becomes finite.[1] Since we cannot operate with infinite quantities, we must say that we have reached the limitations of the theory using a point particle and in order to obtain finite results we must introduce a phenomenological density distribution $\rho(x - z)$ for the particle whose real nature is perhaps outside the domain of classical electrodynamics.

The nature of the mass renormalization term, even if it has been made finite, should be looked either in quantum effects or in general relativistic effects. Quantum electrodynamics, however, has not solved the problem, for the mass-renormalization terms there are still infinite.

Now we must justify Eq. 5.76. In order not to operate with infinite quantities, we write Eq. 5.74 with $\rho(x - z)$ instead of $\delta(x - z)$. It is actually a little simpler to go back to the Lagrangian Eq. 5.62. The mass renormalization term is given by

$$L = \frac{e}{2c} \int ds \dot{z}^{\mu} \rho(x - z)(A_{\mu}^{\text{ret}} + A_{\mu}^{\text{adv}}) \tag{5.79}$$

and

$$\tfrac{1}{2}(A_{\mu}^{\text{ret}} + A_{\mu}^{\text{adv}}) = \frac{e}{2\pi} \int \bar{D}(x - z)\dot{z}_{\mu}\, ds$$

where $\bar{D}$ is the symmetric Green's function

$$\begin{aligned} \bar{D} &= \tfrac{1}{2}(D^{\text{ret}} + D^{\text{adv}}) \\ \bar{D}(x - z) &= \bar{D}(z - x). \end{aligned} \tag{5.80}$$

---

[1] Clearly, mass renormalization is equivalent to throwing away the δm-term completely and identifying m with m_{exp} as before. This shows the phenomenological character of the whole procedure. A different interpretation will be discussed in Chapter VI.

We now proceed in a similar manner as in Section 1, Eq. 5.1.

$$\tfrac{1}{2}(A_\mu^{\text{ret}} + A_\mu^{\text{adv}}) = \bar{A}_\mu = \frac{e}{2\pi}\int ds\,\delta((x-z)^2)\dot{z}_\mu[\theta(x^0 - z^0) + \theta(z^0 - x^0)].$$

We now get contributions from both of the roots of the δ-function $z(s_0)$ and $z(s_1)$ (see Fig. 9).

$$\bar{A}_\mu(x) = \frac{e}{4\pi}\left(\frac{\dot{z}_\mu}{(x-z)^\sigma \dot{z}_\sigma}\bigg|_{s=s_0} + \frac{\dot{z}_\mu}{(x-z)^\sigma \dot{z}_\sigma}\bigg|_{s=s_1}\right).$$

Thus,

$$L = \frac{e^2}{4\pi}\int ds\left[\frac{\dot{z}^\mu(s)\dot{z}_\mu(s_0)}{(x - z(s_0))^\sigma \dot{z}_\sigma(s_0)} + \frac{\dot{z}^\mu(s_1)\dot{z}_\mu(s_1)}{(x - z(s_1))^\sigma \dot{z}_\sigma(s_1)}\right]\rho(x - z(s)). \qquad \textbf{(5.81)}$$

We evaluate the two integrals separately by expanding $z(s)$ around s_0 and s_1, respectively. With $s = s_0 + u$ we have again

$$\dot{z}(s) = \dot{z}(s_0) + u\ddot{z}(s_0) + \frac{u^2}{2}\dddot{z}(s_0) + \cdots$$

$$(x - z(s)) = (x - z(s_0) - u\dot{z}(s_0) + \cdots)$$

The numerator of the first term in Eq. 5.81 is of the form $1 + u^2/2\dddot{z}\cdot\ddot{z}$. If we take terms of the order of u, the numerator is one. For a point particle, the Lagrangian becomes then equal to an infinite constant, we get

$$L = \delta mc \qquad \textbf{(5.82)}$$

with

$$\delta m = -\frac{e^2}{2\pi c}\int du/u,$$

which is precisely of the form of the inertial term of the Lagrangian of a particle.

Dirac[1] has considered a small sphere of radius a in the rest frame of the particle (i.e., a tube along the world line). He calculates the component of the tensor $T^{\mu\nu}$ normal to this surface, $-a^{-1}T^{\mu\nu}n_\nu$, and the energy momentum flow out of this surface,

$$\int\left[\tfrac{1}{2}e^2a^{-1}\ddot{z}_\mu - \frac{e}{c}\dot{z}_\mu(F_{\text{in}}^{\mu\nu} + \tfrac{1}{2}F_{\text{rad}}^{\mu\nu})\right]ds$$

neglecting terms which go to zero with a. Because this expression must depend only on the endpoints the integrand must be a perfect differential, $\dot{b}_\mu$. He chooses $b_\mu = k\dot{z}_\mu$. This is not the only choice but the simplest one.

---

[1] Opus cit. above.

Now the mass renormalization is carried out by putting

$$\tfrac{1}{2}e^2/a - k = m_{\text{exp}}c$$

and one obtains again

$$\begin{aligned} m_{\text{exp}}c\ddot{z}_\mu &= \frac{e}{c}\,\dot{z}_\mu(F^{\mu\nu}_{\text{in}} + \tfrac{1}{2}F^{\mu\nu}_{\text{rad}}) \\ &= K^{\text{ext}}_\mu + K^{\text{self (finite part)}}_\mu \end{aligned}$$

Historically, Lorentz[1] has introduced a three-dimensional extension for the particle as a charged sphere and calculated the radiation reaction as a series in powers of the radius a of the sphere. The term which remains as $a \to 0$ is essentially the finite part which we have evaluated. However, a three-dimensional "size" is not a relativistically invariant concept. Furthermore, such a model for a charged particle would not be stable, since any finite charge distribution would extend under the influence of external forces, unless some additional, say mechanical, forces are introduced to hold the charge distribution together.

It is instructive to develop also an approximative noncovariant derivation of the spatial part of Eq. 5.77 using a space- and time-dependent charge distribution $\rho(\mathbf{x}, t)$. We shall also consider at the end a spherical charge distribution.

We start from the expression

$$\tfrac{1}{2}(A^{\text{ret}}_\mu + A^{\text{adv}}_\mu) = \frac{1}{2c}\int dx'\,\frac{j_\mu(x')}{4\pi|\mathbf{x} - \mathbf{x}'|} \times [\delta(x^0 - x^{0'} - |\mathbf{x} - \mathbf{x}'|) + \delta(x^0 - x^{0'} + |\mathbf{x} - \mathbf{x}'|)].$$

The $dx^{0'}$-integration can be carried out giving

$$\tfrac{1}{2}(A^{\text{ret}}_\mu + A^{\text{adv}}_\mu) = \frac{1}{8\pi c}\int \frac{d^3x'}{|\mathbf{x} - \mathbf{x}'|} \times [j_\mu(\mathbf{x}', x^0 - |\mathbf{x} - \mathbf{x}'|) + j_\mu(\mathbf{x}', x^0 + |\mathbf{x} - \mathbf{x}'|)].$$

We now expand $j_\mu(x^0 \pm |\mathbf{x} - \mathbf{x}'|)$ around the time x^0.

$$\tfrac{1}{2}(A^{\text{ret}}_\mu + A^{\text{adv}}_\mu) = \frac{1}{8\pi c}\int \frac{d^3x'}{|\mathbf{x} - \mathbf{x}'|} \times \left[2j_\mu(\mathbf{x}', x^0) + (\mathbf{x} - \mathbf{x}')^2\,\frac{\partial^2 j_\mu}{\partial x^{02}} + \tfrac{1}{2}|\mathbf{x} - \mathbf{x}'|^4\,\frac{\partial^4 j_\mu}{\partial x^{04}} + \cdots\right].$$

---

[1] H. A. Lorentz, *The Theory of Electrons* (Dover Publ., New York).

The corresponding field is then

$$\tfrac{1}{2}(F_{\mu\nu}^{\mathrm{ret}} + F_{\mu\nu}^{\mathrm{adv}}) = \frac{1}{4\pi c}\int d^3x' \left[\frac{j_{\nu,\mu} - j_{\mu,\nu}}{|\mathbf{x}-\mathbf{x}'|} - \frac{j_\nu \partial_\mu R - j_\mu \partial_\nu R}{|\mathbf{x}-\mathbf{x}'|^2} + \frac{1}{2c^2}\frac{\partial^2}{\partial t^2}(j_\nu \partial_\mu R - j_\mu \partial_\nu R) + \frac{1}{2c^2}\frac{\partial^2}{\partial t^2}(j_{\nu,\mu} - j_{\mu,\nu}) + \cdots\right] \tag{5.83}$$

where we have put $R = |\mathbf{x}-\mathbf{x}'|$.

The electric field is given by

$$\mathbf{E} = -\tfrac{1}{2}(F_{0k}^{\mathrm{ret}} + F_{0k}^{\mathrm{adv}}) = -\frac{1}{4\pi c}\int d^3x' \left(\frac{1}{cR}\frac{\partial j}{\partial t} + \frac{\nabla R}{R^2} j_0 - \frac{\nabla R}{2c^2}\frac{\partial^2 j_0}{\partial t^2} + \cdots\right),$$

or, noting that j^μ; $(c\rho, \rho v)$ we obtain

$$\mathbf{E} = -\frac{1}{4\pi c}\int d^3x' \left(\frac{1}{cR}\mathbf{v}\frac{\partial R}{\partial t} + \frac{1}{cR}\rho\frac{\partial \mathbf{v}}{\partial t} + \frac{\nabla R}{R^2} c\rho - \frac{(\mathbf{x}-\mathbf{x}')}{2cR}\frac{\partial^2 \rho}{\partial t^2} + \cdots\right). \tag{5.84}$$

A similar expression can be written down for the magnetic field.

The contribution of **E** to the self-force can be calculated from the equation

$$\mathbf{F} = \int \rho(\mathbf{x}, x^0) E(\mathbf{x}, x^0)\, d^3x. \tag{5.85}$$

In evaluating this integral we see that due to the equation of continuity

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{v}$$

the first term of Eq. 5.84 does not contribute to **F**. Indeed, if we integrate by parts we get

$$\int d^3x' \frac{\mathbf{v}}{R}\nabla\cdot(\rho\mathbf{v}) = -\mathbf{v}\int d^3x' \rho\mathbf{v}\cdot\nabla\left(\frac{1}{R}\right) = -\mathbf{v}\int d^3x' \rho\mathbf{v}\cdot\frac{(\mathbf{x}-\mathbf{x}')}{R^3};$$

the integrand of Eq. 5.85 corresponding to this term is odd under the exchange of x and x', hence the double integral

$$\iint d^3x\, d^3x' \rho(\mathbf{x})\rho(\mathbf{x}')\mathbf{v}\cdot\frac{(\mathbf{x}-\mathbf{x}')}{R^3} = 0.$$

For the same reason the third term of Eq. 5.84 does not contribute to **F** either. In the last term we make the following steps:

$$\int d^3x \frac{(\mathbf{x}-\mathbf{x}')}{R}\frac{\partial}{\partial t}\frac{\partial\rho}{\partial t} = -\int d^3x' \frac{(\mathbf{x}-\mathbf{x}')}{R}\frac{\partial}{\partial t}\nabla\cdot(\rho\mathbf{v})$$

$$= \frac{\partial}{\partial t}\int d^3x' \rho\mathbf{v}\cdot\nabla\frac{(\mathbf{x}-\mathbf{x}')}{R}$$

$$= \frac{\partial}{\partial t}\int d^3x' \rho\mathbf{v}\cdot\left(\frac{1}{R} - (\mathbf{x}-\mathbf{x}')\frac{(\mathbf{x}-\mathbf{x}')}{R^3}\right)$$

$$= \int d^3x' \left(\frac{\rho\dot{\mathbf{v}}}{R} - \frac{(\mathbf{x}-\mathbf{x}')}{R^3}\rho\dot{\mathbf{v}}\cdot(\mathbf{x}-\mathbf{x}')\right).$$

The first term of this expression combines with the second term of Eq. 5.84 and we finally get the electric force in the form

$$\mathbf{F}^{\mathrm{el}} = -\frac{1}{8\pi c^2}\int d^3x\, d^3x' \frac{\rho(\mathbf{x}, x^0)\rho(\mathbf{x}', x^0)}{R}\left[\dot{\mathbf{v}} + (x - \mathbf{x}')\frac{\dot{\mathbf{v}}\cdot(\mathbf{x}-\mathbf{x}')}{R^2}\right]. \tag{5.86}$$

If the dependence of ρ on $\mathbf{x}$ is spherically symmetric, the second integral is 1/3 of the first integral, therefore

$$\mathbf{F} = -\frac{2}{3}\frac{\dot{v}}{4\pi c^2}\int d^3x\, d^3x' \frac{\rho(\mathbf{x}, t)\rho(\mathbf{x}', t)}{|\mathbf{x}-\mathbf{x}'|}. \tag{5.87}$$

In terms of the electrostatic energy of the charge distribution

$$V_0 = \frac{1}{2}\int \frac{\rho(\mathbf{x}, t)\rho(\mathbf{x}', t)}{|\mathbf{x}-\mathbf{x}'|} d^3x\, d^3x' \tag{5.88}$$

we have finally

$$\mathbf{F} = -\frac{2}{3}\frac{V_0}{2\pi c^2}\ddot{\mathbf{z}} = -\delta m\ddot{z}. \tag{5.89}$$

For a spherical charge distribution of radius a, $V_0 = e^2/a$ hence δm is finite and equal to

$$\delta m = \frac{2}{3}\frac{e^2}{2\pi c^2}\frac{1}{a}.$$

This term becomes infinite as $a \to 0$. There are other terms in the expansion whose contribution to δm vanish as $a \to 0$. We shall come back to the problem of mass renormalization once more at the end of Chapter VI.

6. *Equations of Motion with Radiation Reaction*

Assuming that a mass renormalization has been carried out, the Lorentz-Dirac equation of motion of a charged particle interacting with an external and with its own field is

$$mc\ddot{z}_\mu = \frac{e}{c} F^{\text{ext}}_{\mu\nu}\dot{z}^\nu + \frac{2}{3}\frac{e^2}{4\pi c}(\dddot{z}_\mu + \dot{z}_\mu \ddot{z}^2) \tag{5.90}$$

with the subsidiary condition

$$\dot{z}^\mu \dot{z}_\mu = 1.$$

In the present section we shall consider solutions and interpretation of this equation.

We note that the solution of Eq. 5.90 is the actually observed motion. We can therefore ask first the simple question when a uniform or a uniformly accelerated motion can be observed. To obtain $\dddot{z} = 0$ we must have $F^{\text{ext}}_{\mu\nu} = 0$, but the converse is not true. In the absence of external fields the equation of the motion

$$\ddot{z}_\mu = \tau(\dddot{z}_\mu + \dot{z}_\mu \ddot{z}^2) \tag{5.91}$$

$$\tau = \frac{2}{3}\left(\frac{e^2/4\pi}{mc^2}\right) \tag{5.91$'$}$$

admits solutions other than the physically reasonable one, namely $\ddot{z}_\mu = 0$. Indeed, if we multiply both sides of Eq. 5.91 with $\ddot{z}_\mu$ we obtain

$$\frac{1}{\tau}(\ddot{z}_\mu \ddot{z}^\mu) = \dddot{z}_\mu \ddot{z}^\mu = \frac{1}{2}\frac{d}{ds}(\ddot{z}_\mu \ddot{z}^\mu)$$

which has the solution

$$\ddot{z}^2 = c^2 e^{2s/\tau}. \tag{5.92}$$

Similarly, multiplying both sides of Eq. 5.91 with $\dot{z}_\nu$ and antisymmetrizing we find

$$\frac{1}{\tau}(\ddot{z}_\mu \dot{z}_\nu - \ddot{z}_\nu \dot{z}_\mu) = (\dddot{z}_\mu \dot{z}_\nu - \dddot{z}_\nu \dot{z}_\mu) = \frac{d}{ds}(\ddot{z}_\mu \dot{z}_\nu - \ddot{z}_\nu \dot{z}_\mu)$$

the solution of which is

$$\ddot{z}_\mu \dot{z}_\nu - \ddot{z}_\nu \dot{z}_\mu = K_{\mu\nu}\, e^{s/\tau}. \tag{5.93}$$

Let us choose a frame such that at the proper time $s = 0$,

$$\dot{z}(0) = (1, 0, 0, 0), \qquad \ddot{z}(0) = (0, C, 0, 0),$$

i.e., acceleration in the rest frame at $s = 0$ is in the direction of the x'-axis, then it follows from Eq. 5.93 that

$$K_{10} = -K_{01} = C, \quad \text{all other} \quad K_{\mu\nu} = 0,$$

and we can integrate Eq. 5.93. The result is

$$\begin{aligned} \dot{z}_0 &= \cosh[\tau(e^{s/\tau} - 1)] \\ \dot{z}_1 &= \sinh[\tau(e^{s/\tau} - 1)] \\ \ddot{z}_2 &= 0 \\ \dot{z}_3 &= 0. \end{aligned} \tag{5.94}$$

This solution is clearly unphysical, because the velocity of the free particle tends ultimately to the velocity of light. These are the so-called "*runaway solutions*," or "self accelerated solutions".

The equation of the motion, Eq. 5.90, as it stands, is therefore incomplete and must be supplemented by some additional condition to eliminate the unphysical solutions, if we accept that, within classical limits, it gives us a closed description of the motion.

For the case of the free particle discussed above the solution of the difficulty is immediate. The equation of the motion, Eq. 5.91, is of third order and we need three initial conditions at $s = 0$, namely $z(0)$, $\dot{z}(0)$ and $\ddot{z}(0)$. In the above runaway solution we have assumed an initial acceleration $\ddot{z}(0)^1 = C$ which we were not allowed to do. If we therefore put $C = 0$, we are left with only the physical solution $\ddot{z} = 0$ or $\dot{z} = (1, 0, 0, 0)$ for the particle.

One should have, however, a general physical prescription to choose this third constant of the motion in more complicated cases. In essence the particle should not accelerate over long time intervals more than the external forces are capable of doing it by energy conservation. A general method of incorporating this condition, or equivalently, of choosing the third constant of the motion is obtained, if Eq. 5.90 is transformed into a set of integro-differential equations of motion.

Let us write the first term of Eq. 5.90 as a general time dependent force term, $K_\mu(s)$. It is actually in most cases velocity dependent and therefore unknown. We consider cases where $K_\mu(s)$ considered as a function of s is bounded, even at $s = \pm\infty$. We multiply the equation of the motion

$$\ddot{z}_\mu - \tau\dddot{z}_\mu = K_\mu + \tau\ddot{z}^2\dot{z}_\mu$$

by the integrating factor $e^{-s\tau}$ and obtain the equation

$$-\frac{d}{ds}(\tau\, e^{-s/\tau}\ddot{z}_\mu) = e^{-s/\tau}(K_\mu + \tau\ddot{z}^2\dot{z}_\mu)$$

which we can integrate, so that

$$\tau\, e^{-s/\tau}\ddot{z}_\mu = -\int_0^s ds'\, e^{-s'/\tau}[K_\mu(s') + \tau\ddot{z}^2(s')\dot{z}_\mu(s')] + \tau\ddot{z}_\mu(0). \tag{5.95}$$

If we require that the increase of the acceleration is less than exponential (for the type of bounded forces considered), the left-hand side of Eq. 5.95 will approach zero as $s \to \infty$. Hence we can determine the integration constant to be

$$\ddot{z}_\mu(0) = \frac{1}{\tau}\int_0^\infty ds'\, e^{-s'/\tau}[K_\mu(s') + \tau\ddot{z}^2(s')\dot{z}_\mu(s')].$$

Because of the assumed boundedness of $K_\mu(s')$, this integral will exist. We obtain so the desired integro-differential equation

$$\ddot{z}_\mu(s) = \frac{e^{s/\tau}}{\tau}\int_s^\infty ds'\, e^{-s'/\tau}[K_\mu(s') + \tau\ddot{z}^2(s')\dot{z}_\mu(s')] \qquad \textbf{(5.96)}$$

with τ given by Eq. 5.91′. Clearly every solution of Eq. 5.96 is also a solution of Eq. 5.90; it suffices simply to differentiate this equation if the integral exists. But the converse is not true, because we have required the boundedness of K_μ and, asymptotically, an increase of $\ddot{z}_\mu$ less than the exponential. For example, the runaway solution, Eq. 5.94, for the free particle does not satisfy the integro-differential Eq. 5.96; the integral blows up.

Equation 5.96 is of second order; we need only the initial position and velocity as the third constant of the motion $\ddot{z}_\mu(0)$ has been determined by the asymptotic condition.

We rewrite Eq. 5.96 by a change of variable $s' \to s' + s$

$$\begin{aligned}\ddot{z}_\mu(s) &= \frac{1}{\tau}\int_0^\infty ds\, e^{-s'/\tau}[K_\mu(s'+s) + \tau\ddot{z}^2(s'+s)\dot{z}_\mu(s'+s)]\\ &= \int_0^\infty da\, e^{-a}[K_\mu(s+\tau a) + \tau\ddot{z}^2(s+\tau a)\dot{z}_\mu(s+\tau a)] \qquad \textbf{(5.97)}\end{aligned}$$

Although the solution of the integro-differential Eq. 5.96 are uniquely determined by initial position and velocity, this equation is still quite different from the Newton equations. The causality in the usual sense does not hold: not only is the acceleration at time s determined by all the future motion, but also the motion starts before the action of the force. For example, if the force is zero for $s < s_0$, the acceleration of the particle is different from zero for $s < s_0$. But the time constant for these phenomena as given by the exponential factor $e^{-s/\tau}$ is very small. For electron $\tau = 0.62 \times 10^{-22}$ sec. The classical picture surely breaks down in what happens over such small time intervals (think of uncertainty principle). Even if it did not, these phenomena cannot be observed at the present time. At any rate the *microcausality* would no longer be valid, but the macrocausality, the observable causality, would still be valid.

Some explite solutions of Eqs. 5.90 and 5.96 have been obtained recently by Plass.[1] (See also Exercise 18.)

7. *Theory of the Electromagnetic Mass*

Here we present a different interpretation of the equation

$$m + \delta m = m_{\text{exp}} \tag{5.98}$$

which goes back to Thomson, Lorentz, Abraham, Poincaré, etc. It is the hypothesis that the mass or inertia of the electron is entirely due to its own field; and, furthermore, that the momentum and spin of the particle are momentum and spin of the particle's own field. In other words, we could put $m_0 = 0$ and

$$\delta m = m_{\text{exp}}\,. \tag{5.99}$$

The measured mass of the particle is a result of the motion of the initially massless "particle" in an external field. Although this idea appears to be very attractive it is not possible, at the present time, to build a complete theory on this basis. Certainly the quantum effects must be taken into account. But even within the framework of quantum theories the nature of the mass of the particles remains unexplained. In the early days of relativistic field theories, one used the concept of "bare mass" and "renormalized mass," but the renormalization terms are still infinite in quantum electrodynamics; these concepts seem to be inapplicable to nuclear particles.

The hypothesis of the electromagnetic mass take a simple form for a particle at rest and states that the total relativistic energy (mass) of the particle is equal to the total energy of the fields associated with the particle. The latter quantity for the electromagnetic field is given by

$$\int T^{00}\, d^3x = \frac{1}{2}\int E^2\, d^3x = \frac{e^2}{2(4\pi)^2}\int \frac{d^3x}{r^4} = \frac{e^2}{8\pi}\int_0^\infty \frac{dr}{r^2}$$

and is infinite due to the lower limit of the integral. If we assume a spherical extension for the particle of radius a and equate the total energy to the rest energy mc^2 we find

$$e^2 \int_a^\infty \frac{dr}{r^2} = \frac{e^2}{a} = mc^2 \tag{5.100}$$

or

$$a = \frac{e^2}{mc^2} = 2.8 \cdot 10^{-13}\ \text{cm}.$$

---

[1] G. N. Plass, *Rev. Mod. Phys.*, *33*, (1962), 37.

This quantity is the classical electric radius of the electron which also arises in the calculation of Thomson or Compton cross-sections.[1]

This kind of argument has a certain amount of validity. If we apply the same reasoning to other fields, say the mesonic field of a static neutron[2] which is of the form described by the Yukawa potential

$$V = g \frac{e^{-\mu r}}{4\pi r}$$

where μ is the mass of the meson and g is the analogue of electric charge e, the nucleonic charge, or the so-called coupling constant of the mesonic field to nucleons. We can now compare the masses M and m of a nucleon and an electron, respectively, with the corresponding field energies.[3]

$$\frac{M}{m} = \lim_{a\to 0} \frac{\int_a^\infty T^{00}_{\text{meson}} 4\pi r^2 \, dr}{\int_a^\infty T^{00}_{\text{em}} 4\pi r^2 \, dr}. \tag{5.101}$$

It is interesting that the limit on the right exists. We find

$$\frac{M}{m} = \lim_{a\to 0} - \frac{\int_a^\infty \left[\left(\frac{dV}{dr}\right)^2 + \mu^2 V\right] r^2 \, dr}{\int_a^\infty (e^2/r^2) \, dr}$$

$$= \frac{g^2}{e^2} \lim_{a\to 0} [(\mu a + 1) e^{-2\mu a}] = \frac{g^2}{e^2}. \tag{5.102}$$

If we equate this ratio to the experimental number 1836 we find

$$\frac{g^2}{\hbar c} \cong 13.4. \tag{5.103}$$

The value of $g^2/\hbar c$ obtained in meson physics is very close to this number, ≈15.

A difficulty encountered in the electromagnetic mass hypothesis which has been frequently discussed concerns the transformation properties of the energy and momentum of the field of the particle. The energy and momentum of the field must transform like a four-vector if they have to represent the energy and momentum of the particle.

---

[1] See Exercise 11.

[2] Or of a proton if we neglect the electric field compared to the strong mesonic field.

[3] G. Stephenson, *Nuovo Cim.*, 5, (1957), 1009.

Consider the frame of reference in which the particle is at rest. In this frame $T_{\mu\nu}$ is diagonal, the energy density is

$$T_{00} = \tfrac{1}{2}E^2, \tag{5.104}$$

the momentum density $G = 0$ and

$$T_{11} = (E_1^2 - \tfrac{1}{2}E^2). \tag{5.104$'$}$$

Let us transform into a moving frame x' with

$$\begin{aligned} x^0 &= \frac{1}{\gamma}(x^{0'} + \beta x^{1'}) \\ x^1 &= \frac{1}{\gamma}(x^{1'} + \beta x^{0'}) \\ x^{2'} &= x^2 \\ x^{3'} &= x^3. \end{aligned} \tag{5.105}$$

The energy momentum tensor transform as

$$T'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^\beta}{\partial x'^\nu} T_{\alpha\beta} \tag{5.106}$$

or in component form

$$\begin{aligned} T'_{00} &= \frac{1}{\gamma^2}(T_{00} + \beta^2 T_{11}) \\ T'_{10} &= \frac{\beta}{\gamma^2}(T_{00} + T_{11}) \\ T_{20} &= T_{30} = 0. \end{aligned}$$

The total energy is the integral of T_{00}, $\int T_{00}\, d^3x$. Now the volume element d^3x also transforms

$$d^3x' = \gamma\, d^3x.$$

The energy momentum in the moving frame can be defined as

$$G'_\mu = \int T'_{\mu 0}\, d^3x'. \tag{5.107}$$

Now a covariant definition of the energy momentum

$$G_\mu = \int T_{\mu\nu}\, d\sigma^\nu \tag{5.108}$$

implies necessarily that G_μ is a four-vector. If in addition we have

$$T_{\mu\nu;}{}^{\nu} = 0, \tag{5.109}$$

i.e., free fields, $\int T_{\mu\nu}\, d\sigma^\nu$ is independent of the space-like surface. We can, therefore, take the space-like surfaces at $t = 0$ and $t' = 0$ so that the covariant definition agrees with Eq. 5.107. However, at the position of the particle the condition, Eq. 5.109, is not satisfied. Using Eq. 5.105, we obtain from Eq. 5.107

$$\begin{aligned} G'_0 &= \frac{1}{\gamma}\int (T_{00} + \beta^2 T_{11})\, d^3x \\ G'_1 &= \frac{\beta}{\gamma}\int (T_{00} + T_{11})\, d^3x. \end{aligned} \tag{5.110}$$

Using the value of the electrostatic energy of the field in the rest frame

$$G_0 = \int T_{00}\, d^3x = \tfrac{1}{2}\int E^2\, d^3x = U_0$$

we get

$$G_1 = \int T_{11}\, d^3x = -\tfrac{1}{3}U_0\,.$$

Hence, in the moving frame

$$\begin{aligned} G'_0 &= U_0\,\frac{3+\beta^2}{3(1-\beta^2)} \\ G'_1 &= \tfrac{4}{3}U_0\,\frac{\beta}{\gamma} \end{aligned} \tag{5.111}$$

which do not agree with the transformation of a four-vector. There is no real difficulty or paradox here, because as we have mentioned above, the assumption, Eq. 5.109 does not hold at the position of the particle and, consequently, the covariant definition Eq. 5.108 is not surface independent and does not reduce to Eq. 5.107.

Let us take the definition $G_\mu = \int T_{\mu\nu}\, d\sigma^\nu$. In order to obtain $G_\mu = mv_\mu$ we must choose

$$d\sigma = n^\nu\, d\sigma = \frac{1}{c}\, u^\nu\, d\sigma,$$

i.e., a surface normal to the direction of motion. We have then

$$\begin{aligned} G'_0 &= \frac{1}{c}\int T'_{0\nu}u^\nu\, d\sigma = \int T'_{00}\,\frac{1}{\gamma}\, d\sigma + \int T'_{0k}\,\frac{\mathbf{v}}{\gamma}\, d\sigma \\ G'_1 &= \frac{1}{c}\int T'_{1\nu}u^\nu\, d\sigma = \frac{1}{\gamma}\int T'_{10}\, d\sigma + \frac{1}{\gamma}\int T'_{1k}v\, d\sigma \end{aligned} \tag{5.112}$$

so that the correct definition has additional terms which restore the correct transformation property of the energy momentum four-vector G_μ, as it must be because G_μ is a four-vector by its very definition.

There is no such problem with the transformation of the total charge. Let

$$Q = \int j^0\, d^3x \tag{5.113}$$

be the total charge in the rest frame. In the moving frame

$$j^{0'} = \frac{1}{\gamma} j^0 \qquad (\text{since } \mathbf{j} = 0)$$

and again $d^3x' = \gamma\, d^3x$. Hence

$$Q' = \int j^{0'}\, d^3x' = \int j^0\, d^3x = Q.$$

Of course, the covariant definition

$$Q = \int d\sigma_\mu j^\mu \tag{5.114}$$

with the equation of continuity

$$j^\mu_{,\mu} = 0 \tag{5.115}$$

implies that Q is independent of the space-like surfaces and one can choose surfaces at $t = 0$ and $t' = 0$.

8. *Further Developments and Exercises*

(1) Derive the Lienard-Wiechert Potential using the advanced Green's function.

(2) Calculate the field of a uniformly moving charged particle using Lienard-Wiechert potentials and compare the results with Exercise 5 in Chapter III.

(3) Evaluate the potential of a nuclear particle emitting a classical scalar field.

(4) Show that for a radiation field $F_{\mu\nu}F_{\sigma\rho}$ and $F_{\mu\nu}(x-z)_\sigma$ are symmetric in $(\mu\sigma)$ and $(\nu\sigma)$, respectively.

(5) Prove Eqs. 5.18 and 5.19.

(6) Compare the definition of the radiation field for a single particle as that part of the field which depends on the acceleration and which behaves as $1/R$ at infinity, $F_{\mu\nu}(\ddot{z})$, with the covariant definition, Eq. 5.67.

(7) Show that the field of a uniformly moving particle approaches a null field in the limit of extreme high velocities.

(8) Calculate the energy radiated per unit time and the radiation field of a harmonic oscillator (in the limit $v/c \ll 1$). How long will it take for such a classical oscillator to collapse due to the loss of energy?

(9) *Angular distribution of radiation.* From the discussion preceding Eq. 5.13 we see that for slowly moving charges the angular distribution

of the radiated energy ($1/R$ field) is proportional to $\sin^2\theta$, where θ is the angle between the acceleration $\dot{\boldsymbol{\beta}}$ and the vector $\mathbf{n}$ of the point where one measures the radiation. For fast moving particles we go back to Eq. 5.11. The angular distribution is proportional to

$$\frac{\{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]\}^2}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})^5}$$

or to

$$\sin^2 \theta/(1 - \rho \cos \theta)^5$$

which is peaked forward in the direction θ_{max} with

$$\cos \theta_{\text{max}} = \frac{1}{3\beta}(1 + 15\beta^2 - 1).$$

The greater is the velocity the smaller is θ_{max}.

(10) *Frequency spectrum of the radiation.* If we examine the form of $E(\ddot{z})$ Eq. 5.11, taking, for example, a monochromatic oscillator $v \sim \cos \omega_0 t$ and $\dot{v} \sim \sin \omega_0 t$, we see that due to terms of the form $v\dot{v}$, etc., the radiation is no longer monochromatic.

Also, because the radiation is peaked in the direction of motion a detector will receive a pulse during a time Δt so that the radiation will contain frequencies up to $1/\Delta t$.

If we write the energy radiated per unit time and solid angle in the form $|I(t)|^2$ and if $\tilde{I}(\omega)$ is the Fourier transform of $I(t)$, then by Parseval theorem

$$\int |I(t)|^2 \, dt = \int |\tilde{I}(\omega)|^2 \, d\omega.$$

Hence we can express the total energy radiated over all times by an integral over the frequencies. Hence $\tilde{I}(\omega)$ gives the frequency spectrum of the radiation.

(11) *Electron-photon scattering; classical limit.* If the momentum transfer between the photon and electron is small we can treat the motion of the electron nonrelativistically. For photons we use a plane wave $\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$. The electric field imparts to the electron a velocity

$$\dot{\mathbf{v}} = \frac{e}{m} \mathbf{E}$$

so that the energy radiated by the electron in the solid angle $d\Omega$ and per unit time

$$\frac{dP^0}{dt\,d\Omega} = \frac{e^2}{(4\pi)^2 c^3}\,\dot{v}^2 \sin^2\theta = \frac{1}{2}\,\frac{c}{(4\pi)^2}\,|E_0|^2 \left(\frac{e^2}{mc^2}\right)^2 \sin^2\theta.$$

The differential cross-section, $\sigma(\theta)$, which is the ratio of the energy radiated per unit time and solid angle to the incident energy per unit and time, is then given by

$$\sigma(\theta) = \left(\frac{e^2/4\pi}{mc^2}\right)^2 \sin^2\theta.$$

The total cross-section,

$$\sigma_T = \frac{8\pi}{3}\left(\frac{e^2/4\pi}{mc^2}\right)^2 = \frac{8\pi}{3}\,r_0^2$$

is called the *Thompson cross-section.* The quantum mechanical treatment of this Compton effect is given by the Klein-Nishina formula which agrees with the above classical result only at very small energies and angles.

(12) *Bremsstrahlung.* This is the process of emission of photons or electromagnetic radiation when charged particles in the Coulomb field of other particles change their acceleration. If we can estimate the acceleration of the particle we can use our radiation formula to calculate the so-called Bremsstrahlung.

(13) *Multipole radiation.* So far we have looked at the radiation field in Cartesian coordinates. It is also useful to introduce polar coordinates r, θ, ϕ. Instead of linear momenta k and a superposition of plane waves e^{ikx} one obtains angular momentum variables l and m and *a* superposition of spherical harmonics $Y_{lm}(\theta, \phi)$, or multipole fields.

We recall the well-known multipole expansion of an electrostatic field

$$\Phi(x) = \int \frac{\rho(\mathbf{x}')\,d\mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|}$$

obtained by using the identities

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}}\,P_l(\cos\alpha), \qquad r > r'$$

where a is the angle between $\mathbf{x}$ and $\mathbf{x}'$, and

$$P_l(\cos\alpha) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi)$$

where (θ, ϕ) and (θ', ϕ') are the angles of $\mathbf{x}$ and $\mathbf{x}'$. The result is

$$\Phi(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{4\pi}{2l+1} C_{lm} \frac{Y_{lm}(\theta, \phi)}{r^{l+1}}$$

where the coefficients C_{lm} (multipole moments) are given in terms of the charge distribution ρ:

$$C_{lm} = \int Y_{lm}^*(\theta', \phi') r'^l \rho(\mathbf{x}')\, d\mathbf{x}'.$$

The multipole moments C_{lm} are related to the components of the moments of the charge distribution in Cartesian coordinates, for we have also the expansion

$$\Phi(\mathbf{x}) = \frac{1}{r}\int \rho(\mathbf{x}')\, d\mathbf{x}' + \frac{\mathbf{p}\cdot\mathbf{x}}{r^3} + \frac{1}{2}\sum_{ij} Q_{ij} \frac{x_i x_j}{r^5} + \cdots$$

where $\mathbf{p}$ is the *electric dipole moment*

$$\mathbf{p} = \int \mathbf{x}'\rho(\mathbf{x}')\, d\mathbf{x}'$$

and Q_{ij} are the *quadrupole moments*

$$Q_{ij} = \int (3x_i' x_j' - r'^2\, \delta_{ij})\rho(\mathbf{x}')\, d^3\mathbf{x}'.$$

Now we consider the radiation field of localized oscillatory charges. We start from

$$A^\mu(x) = A^\mu_{\text{out}} + \frac{1}{4\pi c}\int \frac{j^\mu(x')\, dx'}{|\mathbf{x} - \mathbf{x}'|}\, \delta(x^0 - x^{0'} + |\mathbf{x} - \mathbf{x}'|).$$

Assuming oscillatory charges $\mathbf{j}(x) = J_0(\mathbf{x})\, e^{-it\omega}$ and a gauge $(0, \mathbf{A})$, we have

$$\mathbf{A}(\mathbf{x}) = \frac{1}{4\pi c}\int J_0(\mathbf{x}') \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|}\, d\mathbf{x}', \qquad k = \frac{\omega}{c}.$$

At large distances from the charge distribution we can expand

$$|\mathbf{x} - \mathbf{x}'| \cong r - \mathbf{n}\cdot\mathbf{x}'$$

and

$$\frac{e^{-ik\mathbf{n}\cdot\mathbf{x}'}}{\left(1 - \dfrac{\mathbf{n}\cdot\mathbf{x}'}{r}\right)} = 1 + \left(\frac{1}{r} - ik\right)(\mathbf{n}\cdot\mathbf{x}) + \frac{1}{2}\left(\frac{2}{r^2} - \frac{2ik}{r} - k^2\right)(\mathbf{n}\cdot\mathbf{x})^2 + \cdots$$

The contribution of the first term in this expansion gives the electric dipole field, the second term the magnetic dipole and the electric quadrupole fields, and so on. The vector potential for the *dipole field*, for example, is

$$\mathbf{A}(x)_{\text{dipole}} = \frac{e^{ikr}}{4\pi cr}\int J_0(x')\, d^3x = -ik\mathbf{p}\frac{e^{ikr}}{r}$$

where we have used a partial integration and the equation of continuity $\nabla\cdot\mathbf{J}_0 = i\omega\rho$. Thus

$$\mathbf{B}_{\text{dipole}} = \nabla\times\mathbf{A}_{\text{dipole}} = k^2(\mathbf{n}\times\mathbf{p})\frac{e^{ikr}}{r}\left(1-\frac{1}{ikr}\right)$$

$$\mathbf{E}_{\text{dipole}} = \frac{i}{k}\nabla\times\mathbf{B} = k^2(\mathbf{n}\times\mathbf{p})\times\mathbf{n}\frac{e^{ikr}}{r} + [3\mathbf{n}(\mathbf{n}\cdot\mathbf{p})-\mathbf{p}]\left(\frac{1}{r^3}-\frac{ik}{r^2}\right)e^{ikr},$$

outside the sources. At very large distances from the source

$$\mathbf{B}_{\text{dipole}} = k^2(\mathbf{n}\times\mathbf{p})\frac{e^{ikr}}{r}, \qquad \mathbf{E}_{\text{dipole}} = \mathbf{B}_{\text{dipole}}\times\mathbf{n}$$

as it should be. It is then easy to see that the total energy radiated per unit time in the form of the dipole radiation is proportional to the absolute value square of the dipole moment $\mathbf{p}$.

Finally we discuss the separation of the radiation field in angular momentum variables.

The method is similar to the electrostatic case discussed above. We expand the Green's function in terms of the spherical harmonics

$$\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|} = ik\sum_{l=0}^{\infty} j_l(kr')h_l^{(1)}(kr)\sum_{m=-l}^{l} Y^*_{lm}(\theta'\phi')Y_{lm}(\theta,\phi), \quad r>r'$$

where j_l and $h_l^{(1)}$ are the spherical Bessel and Hankel functions respectively. Hence

$$\mathbf{A}(x) = \frac{1}{c}\sum_{l=0}^{\infty} h_l^{(1)}(kr)\sum_{m=-l}^{l}\mathbf{C}_{lm}Y_{lm}(\theta,\phi)$$

with

$$\mathbf{C}_{lm} = \int \mathbf{J}_0(\mathbf{x}')j_l(kr')Y^*_{lm}(\theta'\phi')\,dx'.$$

Once the multipole moments C_{lm} are known from the current distribution $\mathbf{J}_0(\mathbf{x}')$ we can evaluate $\mathbf{A}(\mathbf{x})$ and then the fields by

$$\mathbf{B} = \nabla\times\mathbf{A}, \qquad \mathbf{E} = \frac{i}{k}\nabla\times\mathbf{B}.$$

Note: One can also use directly the equations satisfied by $\mathbf{E}$ and $\mathbf{B}$ in the source free zone

$$(\nabla^2+k^2)\mathbf{B} = 0, \qquad \nabla\cdot\mathbf{B} = 0, \qquad E = \frac{i}{k}\nabla\times\mathbf{B}$$

or, alternatively,

$$(\nabla^2 + k^2)\mathbf{E} = 0, \qquad \nabla \cdot \mathbf{E} = 0, \qquad \mathbf{B} = \frac{-i}{k} \nabla \times \mathbf{E},$$

and solve the Helmholz wave equation $(\nabla^2 + k^2)\mathbf{E} = 0$ in polar coordinates. The result is the same.

A particular solution with fixed l and m is called a *multipole field*; the general form of the radiation is thus obtained as a superposition of multipole fields.

We leave the details and angular distribution of multipole radiation as an exercise.

(14) Discuss the field of a uniformly moving particle in a material media (Cerenkov Radiation).

(15) *Scattering from a bound electron and Rayleigh scattering.* Consider for simplicity a one-dimensional harmonically bound charged particle interacting with an external plane electromagnetic field. The equation of the motion including the radiation reaction is

$$m\ddot{x} = -m\omega_0^2 x + \frac{2}{3}\frac{e^2/4\pi}{c^3}\dddot{x} + eE_0 \cos \omega t$$

where ω_0 is the angular frequency of the oscillator and ω that of the external field. We assume a physically acceptable steady state solution of this equation of the form $x \sim \cos(\omega t + a)$, so that $\ddot{x} = -\omega^2 x$. Hence

$$m\ddot{x} = -m\omega_0^2 x - \frac{2}{3}\frac{e^2/4\pi}{c^3}\omega^2 \dot{x} + eE_0 \cos \omega t$$

The solution of this equation can be written as[1]

$$x(t) = \mathrm{Re}\left(\frac{eE_0/m}{\omega_0^2 - \omega^2 - i\Gamma^2\omega^2}\, e^{-i\omega t}\right)$$

where

$$\Gamma = \frac{2}{3}\frac{e^2/4\pi}{c^3}.$$

---

[1] Write, for example, the equation of the motion in the complex form

$$mz = -m\omega_0{}^2 z + \frac{2}{3}\frac{e^2/4\pi}{c^3}\omega^2 \dot{z} + eE_0\, e^{-i\omega t}$$

and use the

$$z = \text{const}\; e^{-i\omega t}$$

The average acceleration square is given by

$$\ddot{x}^2 = \frac{1}{2}\frac{e^2}{m^2}E_0^2\frac{\omega^4}{(\omega_0^2-\omega^2)^2+\Gamma^2\omega^2}$$

and the radiated energy per unit time

$$\frac{dP^0}{dt\,d\Omega} = \frac{2}{3}\frac{e^2/4\pi}{c^3}\frac{1}{2}\frac{e^2}{m^2}E_0^2\frac{\omega^4}{(\omega_0^2-\omega^2)^2+\Gamma^2\omega^2}.$$

Again if we divide by the initial flux the cross-section for the process is given by

$$\sigma = \sigma_T\frac{\omega^4}{(\omega^2-\omega_0^2)^2+\Gamma^2\omega^2}.$$

In particular for free electrons and small A/ω we get back the Thompson cross-section σ_T (see Exercise 11). But otherwise the cross-section shows a *resonance* at $\omega = \omega_0$. The mechanism for the resonance is the "excitation" of the electron and increase its radiation.

In the region $\omega < \omega_0$ and $\Gamma < \omega_0$ (strongly bounded electron) we get the *Rayleigh Scattering* formula

$$\sigma_R = \sigma_T\frac{\omega^4}{\omega_0^4},$$

the cross-section increases very rapidly with the frequency.

(16) Discuss the nonrelativistic limit of the equation of the motion of a charged particle with radiation reaction, and the nonrelativistic limit of the integro–differential Eq. 5.96. Show that the unique solution of

$$\mathbf{z}(t) = \frac{e^{t/\tau}}{\tau}\int_t^\infty \mathbf{F}(t')\,e^{-t'/\tau}\,dt'$$

for given $F(t)$ is

$$\mathbf{z}(t) = \mathbf{z}(0) + \dot{z}(0)t + \int_0^t dt'\int_0^{t'} dt''\int_0^\infty F(t''+\tau a)\,e^{-a}\,da.$$

(Compare Eq. 5.97.)

(17) *Damped harmonic oscillator.* (Same problem as in Exercise 15.) The most general solution of the equation of the motion for a damped harmonic oscillator, but without external field,

$$\tau\dddot{x} - \ddot{x} - \omega_0^2 x = 0,$$

can be written as

$$x = \sum_{k=1}^{3} C_k \, e^{p_k t}$$

where the p_k's are the roots of the characteristic equation

$$\tau p^3 - p^2 - \omega_0^2 = 0.$$

This root equation has one positive real and two complex conjugate roots with negative real parts. The real root p_1

$$p_1 = (A_+ + A_-)^{-1}; \qquad A_\pm = \left[\frac{\tau}{2\omega_0^2} \pm \left(\frac{\tau}{2\omega_0^2}\right)^2 + (27\omega_0^2)^{-1}\right]^{\frac{1}{3}}$$

leads to a runaway solution. But if we take $C_1 = 0$, the physical solution is

$$x(t) = \frac{\dot{x}(0)}{\omega} \, e^{-t/T} \sin \omega t$$

where

$$\frac{1}{T} = \frac{\omega_0^2}{2}(A_+ + A_-)^2, \qquad \omega = \frac{3\omega_0^2}{2\tau}(A_+^2 - A_-^2).$$

Obtain this solution from the integro-differential equation.

(18) Obtain the exact solution of the one-dimension relativistic equation of the motion with radiation reaction in terms of the external force which is assumed to be independent of velocities. (See G. N. Plass, *Rer. Mod. Phys. 33*, (1962) 37.)

(19) *Passage to quantum theory.* The discussion of the radiation field in Section 2, in particular the energy Eq. 5.35, is a most natural and direct starting point of introducing field quantization as well as ordinary quantum mechanics.

Let us write Eq. 5.35 for a finite volume V in the form

$$P^0 = H = 2V \sum_{\mathbf{k}} k^2 \mathbf{a}(k) \cdot \mathbf{a}^*(k) = \sum_{\mathbf{k}} E_k.$$

The momentum of the field becomes

$$\mathbf{G} = \sum_{\mathbf{k}} (E_k/c)\mathbf{k}.$$

If we introduce the new coordinates

$$\mathbf{Q}_k = (V/c^2)[\mathbf{a}(k) + \mathbf{a}^*(k)]$$

and

$$\mathbf{Q}_k = \mathbf{P}_k = -i\omega(V/c^2)[\mathbf{a}(k) - \mathbf{a}^*(k)]; \qquad \omega = ck$$

we get

$$H = \frac{1}{2} \sum_{\mathbf{k}} (\mathbf{P}_k^2 + \omega^2 \mathbf{Q}_k^2)$$

and the correct Hamilton's equations with this Hamiltonian. Thus the radiation field in a box is a sum of a countable infinite number of oscillators (normal modes) each characterized by a frequency $\omega(k) = ck$, $k = |\mathbf{k}|$. According to Planck's hypothesis each normal mode has a definite energy $\hbar\omega$. But the number of oscillators is a random variable. For example, in the case of thermal equilibrium, the probability of having n normal modes of energy $\hbar\omega$ is $\exp(-n\hbar\omega/kT)$ and one obtains easily the Planck's law of the black body radiation. (Derive this formula.) The alternate and more common interpretation that each oscillator has energies $n\omega\hbar (n = 0, 1, 2, \cdots)$ with probability $\exp(-n\hbar\omega/kT)$ is not correct from the point of view of field quantization.

If we interpret each normal mode of energy $\hbar\omega$ as a relativistic particle of momentum $\hbar k$ and mass zero as far as the energy momentum properties are concerned, as explained in Chapter II, it is natural to ask what corresponds to the electromagnetic field and the field equations $\Box^2 A_\mu = 0$ in the case of relativistic particles of nonzero mass. As discussed briefly in Chapter II, we introduce again the wave vector k and frequency ω by the equations $p = \hbar k$, $E = \hbar\omega$. Then $p^2 - (E/c)^2 = m^2c^2$ implies for these matter waves the equation $k^2 - (\omega/c)^2 = m^2c^2/\hbar^2$ and the corresponding wave equation for which the plane wave decomposition coincides with this relation between k and ω is

$$(\Box^2 + m^2c^2/\hbar^2)\psi = 0,$$

the Klein-Girdon equation, which is the basic equation of quantum theory of massive spinless particles. The nonrelativistic limit of this equation gives the Schrödinger equation. The one particle aspect of the Schrödinger equation has been very much emphasized in textbooks at the expense of its field aspect which in a sense is more fundamental and primary part of quantum theory. Historically quantum theory of fields (at least the free fields) actually preceded the ordinary quantum mechanics.

BIBLIOGRAPHY FOR CHAPTER V

In addition to books listed in Chapter III see:

1. L. A. Lorentz, *The Theory of Electrons* (New York: Dover Publ., 1952); see also his *Collected Papers*, Vol. II, pp. 281, 343.
2. M. Born and E. Wolf, *Principle of Optics* (New York: Pergamon, 1959).
3. J. Hadamard, *Lectures on Cauchy Problem* (New York: Dover Publ., 1953).

4. W. Heitler, *Quantum Theory of Radiation* (New York: Oxford Univ. Press, 1954).
5. L. Rosenfeld, *Theory of Electrons* (Amsterdam: North-Holland, 1951).
6. F. Rohrlich, in *Lectures in Theoretical Physics*, Vol. II, ed. W. E. Brittin and B. W. Downs (New York: Interscience, 1960).

Articles

1. T. Erber, "The Classical Theories of Radiation Reaction," *Fortschritte der Physik*, *9*, (1961), 343.
2. T. Fulton and F. Rohrlich, "Classical Radiation from a Uniformly Accelerated Charge," *Ann. Phys.*, *9*, (1960), 499.
3. F. Rohrlich, "The Equations of Motion of Classical Charges," *Ann. of Physics*, *13*, (1961), 93.
4. A. Schild, "The Radiation Emitted by an Accelerated Point Charge," *Math. Analysis and Appl.*, *1*, (1960), 127.
5. G. N. Plass, "Classical Electrodynamic Equations of Motion with Radiative Reactions," *Rev. Mod. Phys.*, *33*, (1961), 37.
6. F. E. Low, "Correspondence Principle Approach to Radiation Theory," *Amer. Jour. Phys.*, *29*, (1961), 298.
7. P. A. M. Dirac, "Classical Theory of Radiating Electrons," *Proc. Roy. Soc.* (London), *A 167*, (1938), 148.
8. C. J. Eliezer, "The Interactions of Electrons and an Electromagnetic Field," *Revs. Mod. Phys.*, *19*, (1947), 147.
9. J. A. Wheeler and R. P. Feynman, "Interaction with the Absorber as the Mechanism of Radiation," *Revs. Mod. Phys.*, *17*, (1945), 157.
10. F. Rohrlich, "Self Energy and Stability of the Classical Electron," *Amer. Jour. Phys.*, *28*, (1960), 639.
11. R. Haag, "Die Selbstwechselwirkung des Elektrons," *Z. Naturforschung*, *10a*, (1955), 752.

VI.

ACTION AT A DISTANCE ELECTRODYNAMICS

So far, we have considered the field as a physical system on its own right interacting with the particles. When this picture is applied to the case of electromagnetic field interacting with point particles, we found an infinite self-energy, hence the necessity of mass renormalization and the breakdown of microcausality in the motion of charged particles when the radiation reaction is taken into account.

It is also possible to eliminate the concept of field and introduce only action-at-a-distance interactions between the particles. A single particle, in this approach, does not produce a field of its own, hence has no self-energy. The relativistically invariant interactions between the particles is such that they simulate the field between them.

To this class of theories belong the electrodynamics of Wheeler and Feynman which gives a physical interpretation to the finite part of the radiation reaction. Under this interpretation the Dirac prescription of calculating the finite part of self-field (Chapter 5, Sec. 4) and the resulting equation of the motion would be exact.

These questions as well as a covariant treatment of the mass renormalization are discussed in this chapter.

1. *The Action Principle of Fokker-Schwarzschild-Tetrode*

We consider a number of charges e_α of masses m_α, $\alpha = 1, \cdots N$ interacting with each other. The interaction is introduced by means of an action integral W[1]

$$W = \sum_\alpha \int d\tau m_\alpha \dot{z}_\alpha^2 + \sum_\alpha e_\alpha \sum_{\beta \neq \alpha} e_\beta \int d\tau \, d\tau' \dot{z}_\alpha(\tau) \dot{z}_\beta(\tau') \bar{D}(z_\alpha - z_\beta) \qquad \mathbf{(6.1)}$$

where τ and τ' are invariant times for the two particles and $\bar{D}(x)$ is the symmetric Green's function both in z_α and z_β and in past and future.

$$\begin{aligned} \bar{D}(x) &= \tfrac{1}{2}(D^{\text{ret}}(x) + D^{\text{adv}}(x)) \\ \bar{D}(z_\alpha - z_\beta) &= \bar{D}(z_\beta - z_\alpha). \end{aligned} \qquad \mathbf{(6.2)}$$

The only dynamical coordinates in the action W are the particle coordinates $z_\alpha, z_\beta, \cdots$. The second term of W is the sum of all possible two-body interactions which are symmetric in α and β as it should be. This requirement of symmetry fixes the Green's function employed up to a factor.

The Euler Lagrange equations for the α-th particle are

$$m_\alpha \ddot{z}_\alpha^\mu(s) = e_\alpha \dot{z}_\alpha^\nu(s) \sum_{\beta \neq \alpha} \int ds' \, e_\beta \left[\dot{z}_{\beta\nu}(s') \frac{\partial}{\partial z_{\alpha\mu}} - \dot{z}_\beta^\mu(s') \frac{\partial}{\partial z_\alpha^\nu} \right] \bar{D}(z_\alpha - z_\beta) \qquad \mathbf{(6.3)}$$

If we consider the right-hand side of Eq. 6.3 as a force, K^μ, we see that it satisfies, as it should

$$K^\mu \dot{z}_{\alpha\mu} = 0.$$

It represents a force "propagating" with a velocity c, for the force at a proper time s on the α-particle comes from all other particles z_β at proper time s' such that

$$(z_\alpha(s) - z_\beta(s'))^2 = 0.$$

To see the relation of this equation with that of the field theoretical equation of motion we define formally as the field of the particle β at the point x, the quantities

$$A_\mu^{(\beta)}(x) \equiv e_\beta \int ds \bar{D}(x - z_\beta(s)) \dot{z}_{\beta\mu}(s), \qquad \mathbf{(6.4)}$$

hence the corresponding field is

$$\bar{F}_{\mu\nu}^{(\beta)}(x) = e_\beta \int ds \left(\dot{z}_{\beta\nu}(s) \frac{\partial}{\partial x^\mu} - \dot{z}_{\beta\mu} \frac{\partial}{\partial x^\nu} \right) \bar{D}(x - z_\beta(s)). \qquad \mathbf{(6.5)}$$

---

[1] We take in this chapter, for simplicity, the units such that $c = 1$.
H. Tetrode, *Z. Phys.*, *10*, (1922), 317.
J. Frenkel, *Z. Phys.*, *32*, (1925), 518.
G. N. Lewis, *Nt. Acad. Sci.*, *12*, (1926), 22.

Thus, in the sense of the definition, Eq. 6.4, the right-hand side of the equation of motion, Eq. 6.3, involves the sum of the fields of all particles β at the position of the particle α:

$$m_{\alpha}\ddot{z}^{\mu}_{\alpha}(s) = e_{\alpha}\dot{z}^{\nu}_{\alpha}(s) \sum_{\substack{\beta \\ \beta \neq \alpha}} \bar{F}^{\mu(\beta)}_{\nu}(z_{\alpha}). \tag{6.6}$$

If we define a field

$$\bar{F}^{\mu}_{\nu}(z_{\alpha}) = \sum_{\substack{\beta \\ \beta \neq \alpha}} \bar{F}^{\mu(\beta)}_{\nu}(z_{\alpha}) \tag{6.7}$$

which is, according to Eq. 6.2, a sum of retarded and advanced fields. We can put

$$\bar{F}^{\mu(\beta)}_{\nu}(z_{\beta}) = \tfrac{1}{2}(\bar{F}^{\mu(\beta)\text{ret}}_{\nu} + \bar{F}^{\mu(\beta)\text{adv}}_{\nu})$$

also

$$\bar{F}^{\mu}_{\nu}(z_{\alpha}) = \tfrac{1}{2}(\bar{F}^{\mu\,\text{ret}}_{\nu}(z_{\alpha}) + \bar{F}^{\mu\,\text{adv}}_{\nu}(z_{\alpha})).$$

With these definitions the equation of motion

$$m_{\alpha}\ddot{z}^{\mu}_{\alpha} = e_{\alpha}\dot{z}^{\nu}_{\alpha}\bar{F}^{\mu}{}_{\nu} \tag{6.8}$$

is exactly the same as that of a single particle in a given external field $\bar{F}^{\mu}_{\nu}$, given by Eq. 6.7 and Eq. 6.5. Moreover, the field, Eq. 6.7, satisfies Maxwell's equations

$$\begin{aligned} \Box^{2} A^{\mu}(x) &= \sum_{\substack{\beta \\ \beta \neq \alpha}} e_{\beta} \int ds \dot{z}^{\mu}_{\beta}(s) \Box^{2} \bar{D}(x - z_{\beta}) \\ &= \sum_{\substack{\beta \\ \beta \neq \alpha}} e_{\beta} \int ds \dot{z}_{\beta}\, \delta(x - z_{\beta}), \end{aligned} \tag{6.9}$$

if we define the current density of the particle β by

$$\bar{j}_{\beta}^{\mu}(x) = e_{\beta}\, ds \dot{z}^{\mu}_{\beta}\, \delta(x - z_{\beta}),$$

then the right-hand side of Eq. 6.9 is the total current density of all other particles except α, for which it is the effective current density. The subsidiary condition is also satisfied

$$\begin{aligned} A^{\mu}_{,\mu}(x) &= \sum_{\beta} e_{\beta} \int ds \dot{z}^{\mu}_{\beta} \frac{\partial}{\partial x^{\mu}} \bar{D}(x - z_{\beta}) \\ &= -\bar{D}(x - z_{\beta}) \Big|_{s=-\infty}^{+\infty} = 0. \end{aligned} \tag{6.10}$$

Thus we see that the present theory based on an action-at-a-distance action principle is almost equivalent to the Maxwell-Lorentz theory except that the equation of motion, Eq. 6.8, does not contain any term corresponding to radiation reaction. As a matter of fact, there are no radiation effects in the theory, but only interactions of a number of particles. Consequently, there is also no radiation reaction. The field which enters in Eq. 6.8 is, however, of quite a different character than the field produced by a particle in Maxwell-Lorentz theory. Experimentally we observe a radiation. This phenomenon will be interpreted in this theory by taking the detector of the radiation as part of the system. The particles of the detectors (absorber) will be affected by other particles of the system which is interpreted as radiation.

Let us summarize the interpretive statements of the theory.

(1) An accelerated point charge in otherwise particle-free space does not radiate.

(2) The "fields" which act on a given particle arise solely from other particles.

(3) These fields are represented by half the retarded plus half the advanced Lienard-Wiechert solutions of the Maxwell's equations. The law of force so obtained is symmetric with respect to past and future (and with respect to interchange of particles).

(4) To discuss the radiation of a system, we have to assume the existence of charged particles (absorbers) at the boundary of the volume in which the given system is located. The absorbers will be accelerated by the system which is interpreted as the radiation. The absorbers in turn accelerate the system which is the radiative reaction.

To elaborate further the last point, let us write the equation of the motion, Eq. 6.6, in the equivalent form

$$m_\alpha \ddot{z}_\alpha = e_\alpha \dot{z}^\mu_\alpha \left(\sum_{\substack{\beta \\ \alpha \neq \beta}} F^{(\beta)\mathrm{ret}}_{\mu\nu} + \tfrac{1}{2}[F^{(\alpha)\mathrm{ret}}_{\mu\nu} - F^{(\alpha)\mathrm{adv}}_{\mu\nu}] - \frac{1}{2} \sum_{\mathrm{all}\ \beta} [F^{(\beta)\mathrm{ret}}_{\mu\nu} - F^{(\beta)\mathrm{adv}}_{\mu\nu}] \right), \tag{6.11}$$

Here the first term is the contribution of the usual retarded field of all other particles and represents the usual external force in Maxwell-Lorentz theory. The second term is the contribution of the finite part of the self field of the α-particle which, as we have seen, corresponds to the self force à la Dirac. The last term, finally, is zero, if one makes the assumption that all radiation emitted by *all* particles is absorbed by the absorbers:[1] If a source has emitted radiation during a certain time interval which has been completely absorbed, then a very long time afterwards $\sum_\beta F^{(\beta)\mathrm{ret}}_{\mu\nu}$ is zero. The quantity $\sum_\beta F^{(\beta)\mathrm{adv}}_{\mu\nu}$ is also zero at the same time because the source is no longer emitting any radiation. Now because

---

[1] J. A. Wheeler, and R. P. Feynman, *Rev. Mod. Phys.*, *17*, (1945), 157; R. P. Feynman, *Phys. Rev.*, *74*, (1948), 939.

For a comparison of field theory and action-at-distance theory see P. Havas, *Phys. Rev.*, *74*, (1948), 456.

$$\sum_{\beta}\left[F_{\mu\nu}^{(\beta)\mathrm{ret}} - F_{\mu\nu}^{(\beta)\mathrm{adv}}\right]$$

satisfies Maxwell's equations, it vanishes at all times if it vanishes at one time.

For a consistent theory one has to require further that:

(1°) the effect of the absorber on the sources must be independent of the properties of the absorber, and

(2°) this effect must be completely determined by the *motion* of the source itself;

(3°) the absorber must exert on the source a force which is finite and simultaneous with the moment of acceleration and sufficient in magnitude to take away from the source the energy which goes as radiation;

(4°) the absorber produces a field equal to $\frac{1}{2}(F^{\mathrm{ret}} - F^{\mathrm{adv}})$ at the position of the sources. This field combines with the source's $\frac{1}{2}(F^{\mathrm{ret}} + F^{\mathrm{adv}})$ to produce the observed F^{ret}.

With this interpretation the action W gives the correct radiation reaction without an infinite self energy.

2. *Action Principle with Self-Energy*

The self-energy problem is easily discussed with the help of the action W, since it is equivalent to the usual formulation.

We now add a self-energy term in the action W of Eq. 6.1 by writing

$$W' = \sum_{\alpha} m_{\alpha} \int d\tau \dot{z}_{\alpha}^{2} + \frac{1}{2}\sum_{\alpha,\beta} e_{\alpha}e_{\beta} \int ds\, ds' \dot{z}_{\alpha}(s)\dot{z}_{\beta}(s')\bar{D}(z_{\alpha}(s) - z_{\beta}(s')). \quad \mathbf{(6.12)}$$

The action W' differs from W by the term $\alpha = \beta$ which is infinite due to the $\bar{D}$ function. We therefore replace

$$\bar{D} = \frac{1}{2\pi}(\delta(x - x')^{2})$$

by

$$\frac{1}{2\pi}\rho[(x - x')^{2}]$$

where $\rho(x^{2})$ is an invariant function. Instead of Eq. 6.4 we now define a new field

$$\tilde{A}_{\mu}^{(\beta)}(x) = \frac{e_{\beta}}{2\pi}\int ds \dot{z}_{\beta\mu}(s)\rho(x - z_{\beta}(s)) \quad \mathbf{(6.13)}$$

and obtain from Eq. 6.12 for the equation of motion of the α-th particle

$$\begin{aligned} m_\alpha \ddot{z}_{\alpha\nu} &= e_\alpha \dot{z}_\alpha \sum_{\text{all }\beta} \tilde{F}^{(\beta)}_{\mu\nu}(z_\alpha) \\ &= e_\alpha \dot{z}^\mu_\alpha \left[\sum_{\beta\neq\alpha} \tilde{F}^{(\beta)}_{\mu}(z_\alpha) + \tilde{F}^{(\alpha)}_{\mu}(z_\alpha) \right] \qquad (6.14) \\ &= K^{\text{ext}}_\mu + K^{\text{self}}_\mu . \end{aligned}$$

We remark that the field, Eq. 6.13, does not, strictly speaking, obey the Maxwell equations on account of ρ instead of δ-function density:

$$\Box^2 \tilde{A}^{(\beta)}_\mu = \frac{e_\beta}{2\pi} \int ds \dot{z}_{\beta\mu} \Box^2 \rho(x - z_\beta),$$

unless we modify the current density slightly as one should.

The subsidiary condition

$$\begin{aligned} A^\mu_{,\mu} &= \frac{e_\beta}{2\pi} \int ds \frac{dz_{\beta\mu}}{ds} \frac{\partial}{\partial x^\mu} \rho(x - z_\beta) \\ &= -\frac{e_\beta}{2\pi} \int d\rho(x - z_\beta(s)) \end{aligned}$$

is satisfied if ρ falls off rapidly as $s \to \pm\infty$. One can choose $\rho(x^2)$, for example, in the following form

$$\begin{aligned} \rho(x^2) &= \frac{1}{2a^2} \exp\left(-|x^2|/a\right) \quad &\text{for} \quad x^2 > 0 \\ &= 0 &\text{for} \quad x^2 < 0 \end{aligned} \qquad (6.15)$$

where a is a characteristic length, say, of the order of the electron radius $a \approx e^2/mc^2$, so that it behaves like a δ-function at large distances where Maxwell's equations are strictly satisfied.

We can write the equation of motion, Eq. 6.14, in a slightly different form by eliminating the term absorbed by the absorber:

$$m_\alpha \ddot{z}_{\alpha\nu} = e_\alpha \dot{z}^\mu_\alpha \sum_{\text{all }\beta} \tilde{F}^{(\beta)}_{\mu\nu} = e_\alpha \dot{z}^\mu_\alpha \left[\sum_{\text{all }\beta} \bar{\bar{F}}^{\text{ret}}_{\mu\nu} - \frac{1}{2} \sum_{\text{all }\beta} \left(F^{(\beta)\text{ret}}_{\mu\nu} - F^{(\beta)\text{adv}}_{\mu\nu} \right) \right] \qquad (6.16)$$

where we have introduced

$$\bar{\bar{F}}^{\text{ret}} = \tilde{F} + \tfrac{1}{2}(F^{\text{ret}} - F^{\text{adv}}).$$

In Eq. 6.16 the second term on the right-hand side is again zero, hence

$$m_\alpha \ddot{z}_{\alpha\nu} = e_\alpha \dot{z}^\mu_\alpha \sum_{\text{all }\beta} \bar{\bar{F}}^{\text{ret}}_{\mu\nu}, \qquad (6.17)$$

so that $\bar{\bar{F}}^{\text{ret}}$ appears as the effective field. If we separate once more the damping and self-force terms, we get

$$m_\alpha \ddot{z}_{\alpha\nu} = e_\alpha \dot{z}_\alpha^\mu \left[\sum_{\substack{\beta \\ \beta \neq \alpha}} \bar{\bar{F}}_{\mu\nu}^{(\beta)\text{ret}} + \tilde{F}_{\mu\nu}^{(\alpha)} + \tfrac{1}{2}(F_{\mu\nu}^{(\alpha)\text{ret}} - F_{\mu\nu}^{(\alpha)\text{adv}}) \right]$$

$$= K_\nu^{\text{et}} + K_\nu^{\text{self}} + K_\nu^{\text{rad. reaction.}} \tag{6.18}$$

which is the same result as obtained in the previous section.

3. *Mass Renormalization*

We show in this section how the self energy term can be incorporated into the inertia or mass term in a covariant way. The term $\alpha = \beta$ in the action W', Eq. 6.12, with an invariant density ρ has the form

$$\frac{1}{4\pi} e_\alpha^2 \int ds\, ds' \dot{z}_\alpha(s') \rho(z_\alpha(s) - z_\alpha(s')). \tag{6.19}$$

Since ρ is peaked like a δ-function only the values of $z_\alpha(s')$ close to $z_\alpha(s)$ is important, or, if we let $s' = s + u$, only small values of u. We therefore expand $z_\alpha(s + u)$ around s. Then

$$(z_\alpha(s') - z_\alpha(s))^2 = u^2 + O(u^4)$$

$$\dot{z}_\alpha(s') = \dot{z}_\alpha(s) + u\ddot{z}_\alpha(s) + \frac{u^2}{2} \dddot{z}_\alpha(s) + \cdots$$

and Eq. 6.19 becomes

$$\frac{1}{4\pi} e_\alpha^2 \int ds\, du \dot{z}_\alpha(s) \left[\dot{z}_\alpha(s) + u\ddot{z}_\alpha(s) + \frac{u^2}{2} \dddot{z}_\alpha(s) + \cdots \right]$$

$$\times \rho(u^2 + O(u^4)). \tag{6.20}$$

The second term gives a vanishing contribution since $\dot{z}_\alpha \ddot{z}_\alpha = 0$. The first term gives

$$\frac{1}{4\pi} e_\alpha^2 \int ds \int_{-\infty}^{+\infty} du \rho(u^2)$$

or, if we introduce instead of s the arbitrary parameter

$$\frac{1}{4\pi} e_\alpha^2 \int d\tau z_\alpha^2 \int_{-\infty}^{+\infty} du \rho(u^2) = \int d\tau\, \delta m_\alpha \dot{z}_\alpha^2 \tag{6.21}$$

where we have defined

$$\delta m_\alpha = \frac{1}{4\pi} e_\alpha^2 \int_{-\infty}^{+\infty} du \rho(u^2). \tag{6.22}$$

We get thus a self-energy term which has the same form as the first term of W or W'.

For the function ρ given by Eq. 6.15 we find

$$\delta m_\alpha = \frac{e_\alpha^2}{4\pi}\frac{1}{2a^2}\int_{-\infty}^{+\infty} du\, e^{-u/a}$$
$$= \frac{e_\alpha^2}{4\pi}\frac{1}{a}\,.$$

Adopting the hypothesis of the electromagnetic mass and equating $\delta m = m_{\text{exp}}$, we get, with correct units,

$$a = \frac{1}{4\pi}\frac{e^2}{mc^2}$$

as we have anticipated.

The higher order terms in the expansion, Eq. 6.20, give, however, correction to our result that the self energy term is completely inertial. For, example, the third term gives

$$\frac{1}{4\pi}e_\alpha^2\int du\,\frac{u^2}{2}\,\rho(u^2)\int ds\dot{z}_\alpha(s)\dddot{z}_\alpha(s)$$

which, unfortunately, is not of the form of Eq. 6.21. Thus, with a covariant cut-off the problem of mass renormalization on classical level is not completely solved within this approach.

RELATED MATHEMATICAL BOOKS

B. B. Baker and E. T. Copson, *Mathematical Theory of Huygen's Principle*, 2nd Ed. (New York: Oxford Univ. Press, 1950).

R. V. Churchill, *Fourier Series and Boundary Value Problems* (New York: McGraw-Hill, 1941).

R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I and II (New York: Interscience, 1953 and 1962).

B. Friedman, *Principles and Techniques of Applied Mathematics* (New York: John Wiley, 1956).

M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (New York: Cambridge Univ. Press, 1958).

W. Magnus and F. Oberhettinger, *Special Functions of Mathematical Physics* (New York: Chelsea, 1949).

P. M. Morse and H. Feshback, *Methods of Mathematical Physics*, Vol. I and II (New York: McGraw-Hill, 1953).

A. Sommerfeld, *Partial Differential Equations in Physics* (New York: Acad. Press, 1949).

E. J. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd Ed. (Oxford Univ. Press, 1948).

G. N. Watson, *Theory of Bessel Functions*, 2nd Ed. (Cambridge Univ. Press, 1953).

AUTHOR INDEX

AUTHOR INDEX

A

Aharoni, J., 45

B

Bade, W. J., 45
Baker, B. B., 221
Bargmann V., 75
Barut, A. O., 122
Bateman, H., 129, 159
Becker, R., 130
Belinfante, F., 114
Bergmann, P. G., 45
Bland, G. F., 4
Blass, G. A., 41
Bogoliubov, N. N., 130
Bopp, F., 131
Born, M., 211

C

Cartan, E., 26
Cederholm, J. P., 4
Chang, T. S., 131
Churchill, R. V., 221
Coish, H. R., 164
Copson, E. T., 221
Corson, E. M., 45
Courant, R., 87, 221
Cunningham, E., 129

D

De Broglie, L., 53
Dirac, P. A. M., 187, 212
Durbin, R. P., 22

E

Einstein, A., 52
Eliezer, C. J., 212
Erber, T., 212

F

Feshbach, H., 221
Feynman, R., 212, 214
Fock, V., 41, 45
Foldy, L. I., 45
Frenkel, J., 83, 214
Friedman, B., 221
Fronsdal, C., 131
Fulton, T., 212

G

Goldstein, H., 24, 107
Graeub, W., 10

H

Haag, R., 212
Hadamard, J., 211
Halbwachs, F., 83
Havas, P., 104, 214
Havens, L., 4
Havens, W. W., 22
Heitler, W., 212
Hess, S., 131
Hilbert, D., 87, 221

I

Infeld, L., 66
Ivanenko, D., 130
Ives, H. E., 6

J

Jackson, J. D., 130
Jehle, H., 45
Jost, R., 34, 45

K

Kahana, S., 164
Kallman, G., 83
Kennedy, R. J., 6

L

Landau, L., 80, 130
Lewis, G. W., 214
Lichnerowitz, A., 45
Lifshitz, E., 80, 130
Lighthill, H. J., 221
Loar, H. H., 22
Lorentz, H. A., 193, 221
Low, F. E., 212

M

Macfarlane, A. J., 45
Magnus, W., 221
Mason, M., 130
Mazur, P., 127
Michel, L., 15, 75
Mills, R. L., 164
Moller, C., 45
Morse, P. M., 221
Moses, M. E., 131
Mullen, G., 122

N

Nijboer, B. R. A., 127
Noether, E., 130

O

Oberhettinger, F., 221

P

Panofsky, W. K. H., 130
Papapetrou, A., 45
Pauli, W., 3, 45, 130
Penrose, R., 22
Peres, A., 84
Phillips, H., 130
Plass, G. H., 199, 210, 212
Proca, A., 81, 84

Q

Quam, P. M., 83

R

Robertson, H. P., 6
Rohrlich, F., 212
Rosen, N., 84
Rosenfeld, L., 212
Rzewuski, J., 130

S

Sakurai, J., 164
Sauter, F., 84, 130
Schild, A., 41, 45, 212
Schiller, R., 84
Schwinger, J., 164
Sciama, D. W., 131
Shirkov. A., 130
Sokolov, A., 130
Sommerfeld, A., 130, 221
Stephenson, G., 200
Synge, J. L., 45

T

Taylor, J. G., 131
Telegdi, V., 75, 84
Terrell, J., 22
Tetrode, H., 214
Thirring, W., 130
Thorndike, E. M., 6
Tischer, M., 131
Titchmarsh, E. J., 221
Townes, C. H., 4

V

van Vleck, J., 127
von Laue, M., 185

W

Watson, G. N., 221
Weaver, W., 130
Weisskopf, V., 22
Werle, J., 80
Wheeler, J. A., 212, 214
Wightman, A. S., 34, 43, 45
Wigner, E. P., 14, 41
Winogradzki, J., 131
Winston, R., 84
Wolf, E., 211

Y

Yang, C. N., 164

SUBJECT INDEX

SUBJECT INDEX

A

Aberration, 20
Absorber theory, 214
Acceleration four vector, 48
Action at-a-distance electrodynamics, 213
Action principle
  for a particle, 60
  for fields, 99
  for fields and particles, 14
  of Fokker-Schwarzschild-Tetrode, 213
  with self-energy, 215
Addition (relativistic) of velocities, 18
Advanced Green's functions, 155
Ampère's law, 89
Analysis, 38
Angular distribution of radiation, 203
Angular momentum, 57, 129, 163
  total, 110
Appearance of moving objects, 41

B

Basis, 11
  orthogonal, 11, 83
Bremsstrahlung, 205

C

Canonical equations
  for a particle, 68
  for fields, 119
  first order, 179
Canonical transformations, 72
Cauchy problem, 160
Causality requirement, 8
Center of mass frame, 59, 80
Čerenkov radiation, 208
Charge, 88, 112, 202
Charge density, 89
Classical radius of the electron, 199 205
Clock "paradox", 41
Complex conjugate of a matrix, 12
Complex orthogonal group, 43
Compton effect, 205
Conformal invariance, 129
Conservation laws, 105, 163
  in integral form, 105
  in differential form, 112
  for colliding particles, 59
Conserved quantities
  explicit form of, 116, 117
Constants of motion, 64, 105
Coordinate transformations, 108
Coulomb field, 167, 168
Coulomb gauge, 95
Contours of integration, 152
Covariance, 35
Covariant description of spin, 44
Covering group, 43

Curl, 38
Current, 89, 90, 91, 92
Current density four vector, 94, 139
Currents, external, 135

D

D'Alembert's equation, 95, 149, 161
Damped harmonic oscillator, 209
Delta function, 139, 154, 156
Dielectric constant, 91
Doppler effect, 20
Dynamics (relativistic), 47

E

Electric field, 88
Electric moments, 73
Electrodynamics of Lorentz, 92
Electromagnetic fields, 54, 88, 177, 118
  constant, 80, 127
  transverse and longitudinal components of, 177
Electromagnetic mass, 199
Electron-Photon scattering, 204
Energy balance, 184
Energy density, 118
Energy momentum tensor, 107, 109, 111, 128, 163, 201
  canonical unsymmetric 113
  symmetric, 111
Equation of continuity, 89, 92, 136
Equation of motion
  for a particle, 68
  for spin, 75
  integro-differential, 198
  with radiation reaction, 185, 195
Euler-Lagrange equations,
  for a particle, 62
  for fields, 87, 100
Exponential matrix, 32

F

Faraday's law, 89, 127
Fermi coordinates, 83
Field equations, 100
  canonical form of, 119
Field tensor, 168
Fields,
  Dirac, 102
  electromagnetic, 54, 88, 102
  scalar, 102
Fitzgerald contraction, 21
Force density, 137
Four vector, 6, 7
  contravariant, 7
  covariant, 7
  time-like, light-like, space-like, 8
Fourier decomposition, 174, 210

G

Galilean transformations, 4
Gauge
  Coulomb, 95
  Lorentz, 95
  radiation, 95, 177
Gauge fields, 146, 163
Gauge transformations, 94
  of the first kind, 142
  of the second kind, 143
Gauss' theorem, 39, 40, 90
Generalized momenta, 63
Generators
  of invariant transformations, 64
  of variations, 106
Gravitational fields, 55
Green's functions, 148
  advanced, 155
  for Dirac field, 164
  for scalar field, 158
  other invariant, 157
  retarded, 153

H

Hamiltonian, 65, 122
Hamiltonian density, 120
Hamilton-Jacobi theory, 73, 82
Hamilton's equation
  covariant form, 120
  for spin, 76
Hamilton's principle (covariant), 60

Hermitian conjugate, 12
Hypercharge, 148

I

Inertial frame of reference, 3
Integro -Differential equation of motion, 198
Interacting fields and particles, 186
Interactions
  between fields, 142
  of fields with a particle, 138
  of fields with external currents, 135
Internal space, 147
Invariance, 35
Invariants of the electromagnetic field, 95, 171
Isotopic spin space, 147

K

Kennedy-Thorndike experiment, 6
Kepler problem, 80
Klein-Gordon equation, 101, 211
Klein-Nishina formula, 205
Kronecker product, 28

L

Lagrange multipliers, 65
Lagrangian
  containing a subsidiary condition, 122
  density, 87
  involving higher order derivatives, 122
  not invariant under translations, 115
Lagrangian form of equations of motion
  for a field, 99
  for a particle, 60
Lienard-Wiechert potential, 127, 165, 167, 203
Light cone, 8
Linear chain, 86, 121
Little group, 44
Lorentz electron, 193
Lorentz force, 54, 55, 62, 63, 88
Lorentz group, 23
  complex, 34
  full homogeneous, 22
  gauge dependent representations of, 179
  infinitesimal generators of, 32
  inhomogeneous, 22
  restricted, 23
  orthochronous, 23
  representations of, 22
Lorentz space, 6, 8
  complex, 11
Lorentz transformations, 3
  antichronous, 13
  decomposition of, 14, 18
  Eigenvalues of, 14, 41
  homogeneous and inhomogeneous, 6, 13
  infinitesimal, 17
  into a moving frame, 17
  linearity of, 5
  orthochronous, 13
  proper and improper, 12
  the physical basis of, 3

M

Magnetic field, 88
Magnetic moment, 73
Magnetic permeability, 91
Mass renormalization, 190, 219
Matter wave, 53
Maxwell's equations, 90, 94, 136
  in material media, 96, 127
  in spinor form, 97
Maxwell field, 102, 103
Maxwell-Lorentz equations, 93
Michelson-Morley experiment, 4
Micro causality, 198
Microscopic equations of electrodynamics, 92
Minkowski equation, 50, 54, 62
Minkowski force, 54
Momentum density, 118
Momentum four vector, 48
Multipole radiation, 207

N

Neutrino, 53, 80
Nucleon number, 148
Null fields, 170, 172

O

Ohm's law, 92
Orthogonal group, 4, 24, 43

P

Parity, 16, 23
Particles with zero rest mass, 51
Pauli matrices, 24, 25, 43
Photon, 51
Plane reflection or symmetry, 14
Poisson bracket, 72
Polar decomposition, 44
Polarization P., 90
Polarization
  covariant description of, 44
  of photons, 173
Poynting vector, 118, 170
Principle of relativity, 35
Proper time formalism, 47, 48

Q

Quantum electrodynamics, 191
Quantum theory, 53, 210

R

Radiation, 165
  angular distribution of, 203
  energy momentum of, 179
  frequency spectrum of, 204
Radiation field, 169, 170
  canonical formalism, 177
  energy momentum of the, 175
  plane wave decomposition of, 171
  pure, 172
  transverse character, 173
Radiation gauge, 95, 177
Radiation reaction, 184
  equation of motion with, 185, 195
Rayleigh scattering, 208, 209
Relativistic
  dynamics, 47
  field theory, 85
  kinetic energy, 51
Representations
  higher order spinor and tensor, 29
  irreducible and reducible, 29
Rest energy, 51
Rest mass, 48
Rest mass zero, 51
Retarded Green's functions, 153
Retarded time, 168
Runaway solutions, 197

S

Scalar field, 36, 102, 117, 121
  equation of motion in external, 80
  particle in a, 141
Scattering
  electron photon, 204
  experiments, 162
  from a bound electron, 208
  matrix, 162
Schrödinger equation, 211
Schwarz inequality, 10
Self force, 186
  finite part of, 187
Space-like surfaces, 40
Space-time reflection, 16
Spin
  classical, 73
  covariant description of, 44
  precession in field theory, 129
Spinor fields, 36, 38
Spinor mechanics, 81
Spinor space, 27
Spinors, 26
  higher order, 29
Strange particles, 148
Strong interactions, 148
Structure constants, 34
Systems of colliding particles, 58

T

Tensor fields, 36, 37
Tensors, 26

Thomas precession, 82
Thompson cross section, 205
Time dilatation, 21
Time reflection, 16
Total
  angular momentum, 110
  energy, 110
  energy-momentum tensor, 136
Transpose of a matrix, 12
Transverse character of the electromagnetic field, 173
Two-component spinors, 26

U

Unimodular group, 23
  infinitesimal generators of, 32
Unitary unimodular group, 31

V

Variations
  with a fixed boundary, 99
  intrinsic, 112
  involving a change of boundary, 103
Vector fields, 36, 144
Vector potential, 63, 93, 96
Velocity four vector, 48

W

Wave equation, 93, 172
  inhomogeneous, 149, 161
Weysenhoff condition, 79
World line, 47, 139, 158, 166

Y

Yang-Mills fields, 146, 164
Yukawa potential, 141, 200

XAVIER UNIVERSITY LIBRARY
3 1303 00076 3830

(*continued fron*

LIGHT SCATTERING BY SMALL PARTICLES, H. C. van de Hulst. (64228-3) $7.50

SOURCES OF QUANTUM MECHANICS, B.L. Van der Waerden (ed.). (61881-1) $6.95

UNDERSTANDING THERMODYNAMICS, H.C. Van Ness. (63277-6) $3.50

SPAC

THI

TRI ng

VA ng

PE an

F out
n pt.
F Y.
1 ver
2 es,
li ci-
e A.

ELECTRODYNAMICS AND CLASSICAL THEORY OF FIELDS AND PARTICLES

A.O. BARUT

> ". . . we can only hope that more such striking expositions will be written."
> —Bulletin of the American Mathematical Society

Students of theoretical physics interested in the classical underpinnings of modern quantum field theory will welcome this advanced text. It is a systematic, covariant treatment of the classical theories of particle motion, fields, and the interaction of fields and particles. Particular attention is given to the interaction of charged particles with the electromagnetic field.

The treatment throughout the book is relativistic; the author attempts, as much as possible, a coordinate free (or covariant) form of the equations both for particles and the fields. The book opens with an extensive discussion of space-time, Lorentz transformations, Lorentz-group and tensor and spinor fields. This material is essential to the understanding of many branches of theoretical physics, in particular relativistic quantum theory. Chapter II describes various relativistic forms of the fundamental problem of dynamics: describing the trajectories of particles for given external forces. The general dynamical principles to obtain the field equations and the important problems of the conservation laws are discussed in Chapter III.

The second part of the book (Chaps. IV–VI) is devoted to a lucid treatment of the *interactions* of fields and particles. Chap. IV deals with equations of motion and their solutions (the so-called Cauchy problem), focusing on the solution of field equations with Green's Functions using Dirac formalism. The problem of feedback between particles and fields (radiation and radiation reaction) is taken up in Chap. V, as are questions concerning the limitations of classical field theories and classical dynamics. Dr. Barut concludes the book with an excellent exposition of the purely mechanical approach to the problem of the interactions of charged particles—the so-called action-at-a-distance formulation of electrodynamics.

Written on a sophisticated level suited to mature graduate students, this unique and important work includes many discussions of hitherto unpublished original work. Indeed, it is considered by many to be the first comprehensive treatment of relativistic electrodynamics ever written. To further assist the student, Prof. Barut (a highly distinguished theoretical physicist now at the University of Colorado, Boulder) has included problems, alternate proofs, and additional topics at the end of each chapter, where a useful bibliography is also provided. A general bibliography appears at the end of the book.

Unabridged, corrected (1980) republication of the edition published by Macmillan & Co., New York, 1964. Author and Subject Indexes. General and Chapter Bibliographies. 10 figures. 256pp. 5⅜ x 8½. Paperbound.

ISBN 0-486-64038-8 $5.50 in U.S.

Cover
P7-BGI-362